Accurate 3D-reconstruction and -navigation for high-precision minimal-invasive interventions

vom Fachbereich Informatik
der Technischen Universität Darmstadt
genehmigte

DISSERTATION

zur Erlangung des akademischen Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
von

Dipl.-Ing. Wissam El Hakimi

geboren in Le Kef, Tunesien

Referenten der Arbeit:	Prof. Dr.-Ing. Georgios Sakas Technische Universität Darmstadt
	Prof. Dr. techn. Dieter W. Fellner Technische Universität Darmstadt
	Prof. Dr. med. Dr. h. c. Jörg Schipper Heinrich-Heine-Universität Düsseldorf
Tag der Einreichung:	11/12/2015
Tag der mündlichen Prüfung:	03/02/2016

Darmstadt 2016
D 17

Erklärung zur Dissertation

Hiermit versichere ich die vorliegende Dissertation selbständig nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die aus Quellen entnommen wurden, sind als solche kenntlich gemacht. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Darmstadt, den 11/12/2015 Wissam El Hakimi

Abstract

The current lateral skull base surgery is largely invasive since it requires wide exposure and direct visualization of anatomical landmarks to avoid damaging critical structures. A multi-port approach aiming to reduce such invasiveness has been recently investigated. Thereby three canals are drilled from the skull surface to the surgical region of interest: the first canal for the instrument, the second for the endoscope, and the third for material removal or an additional instrument. The transition to minimal invasive approaches in the lateral skull base surgery requires sub-millimeter accuracy and high outcome predictability, which results in high requirements for the image acquisition as well as for the navigation.

Computed tomography (CT) is a non-invasive imaging technique allowing the visualization of the internal patient organs. Planning optimal drill channels based on patient-specific models requires high-accurate three-dimensional (3D) CT images. This thesis focuses on the reconstruction of high quality CT volumes. Therefore, two conventional imaging systems are investigated: spiral CT scanners and C-arm cone-beam CT (CBCT) systems. Spiral CT scanners acquire volumes with typically anisotropic resolution, i.e. the voxel spacing in the slice-selection-direction is larger than the in-the-plane spacing. A new super-resolution reconstruction approach is proposed to recover images with high isotropic resolution from two orthogonal low-resolution CT volumes.

C-arm CBCT systems offers CT-like 3D imaging capabilities while being appropriate for interventional suites. A main drawback of these systems is the commonly encountered CT artifacts due to several limitations in the imaging system, such as the mechanical inaccuracies. This thesis contributes new methods to enhance the CBCT reconstruction quality by addressing two main reconstruction artifacts: the misalignment artifacts caused by mechanical inaccuracies, and the metal-artifacts caused by the presence of metal objects in the scanned region.

CBCT scanners are appropriate for intra-operative image-guided navigation. For instance, they can be used to control the drill process based on intra-operatively acquired 2D fluoroscopic images. For a successful navigation, accurate estimate of C-arm pose relative to the patient anatomy and the associated surgical plan is required. A new algorithm has been developed to fulfill this task with high-precision. The performance of the introduced methods is demonstrated on simulated and real data.

Zusammenfassung

Im Rahmen des DFG-Projekts MUKNO (Multi-Port Knochenchirurgie am Beispiel der Otobasis, FOR 1585), wurde erstmals ein System eines Multi-Port-Zuganges an der Otobasis entwickelt. Es verfolgt das Ziel eines gewebeschonenden minimaltraumatischen Operationsverfahrens. Hierbei werden drei Bohrkanälen zu einem Zielpunkt innerhalb des Felsenbeins gebohrt, die dann synchron für einen Endoskop zur Visualisierung des Zielgebietes sowie für zwei Instrumentenkanäle genutzt werden können.

Ein Überblick über den Prozess des Multi-Port Eingriffs ist in Abbildung 0.1 gegeben: Es wird nach der Diagnose entschieden ob der Patient für eine MUKNO-Operation geeignet ist. Wenn ja, wird am Tag der Operation nach der Allgemeinanästhesie die Grundplatte inklusive Marker am Patienten angebracht. Eine hochauflösende Computertomographie (CT) wird dann aufgenommen und basierend darauf werden die relevanten Risikostrukturen segmentiert und eine hochpräzise Planung der Bohrkanäle erstellt [Bec14]. Danach wird die Positioniereinheit [NPS13] für den Bohrer auf die Grundplatte montiert. Nach der Übertragung der Planung in das lokale Koordinatensystem der Positioniereinheit mit Hilfe der Marker kann mit dem Bohrprozess begonnen. Während der Anlegung der Zugänge wird die Lage des Bohrers punktuell mit einem C-Bogen überprüft. Anschließend wird die Operation im engeren Sinne durchgeführt.

Die Anatomie des Felsenbeins ist äußerst komplex und beinhaltet eine enge Lagebeziehung der sensorischen Organe (Hörschnecke und Gleichgewichtsorgan), Nerven und vitale Gefäße. Um eine genaue Lokalisierung der anatomischen Strukturen zu ermöglichen, müssen die akquirierten Bilddaten mit möglichst hoher isotroper Auflösung aufgenommen werden: Die Präzision der Bildgebung ist entscheidend für die Genauigkeit der Segmentierung, Planung und den eigentlichen Bohrprozess. In dieser Arbeit werden verschiedene Verfahren zur hoch-präzisen 3D Rekonstruktion sowie Navigation für die Multi-Port Otobasischirurgie entwickelt. Dafür werden zwei Ansätze untersucht:

- Super-Resolution Rekonstruktion (SRR) zur Rekonstruktion hochaufgelöster isotroper Voxeldaten aus zwei niedrigaufgelöster CT Volumen.
- Die direkte Aufnahme hochaufgelöster isotroper Voxeldaten mittels eines C-Bogens.

Weiterhin werden Methoden zu Navigation entwickelt, um präzise intra-operative Kontrolle der Instrumentenlage (Soll-Ist-Vergleich) zu ermöglichen.

Super-Resolution Rekonstruktion (SRR)

Ein Spiral CT hat den Nachteil, anisotrope Bilddaten zu akquirieren: i.A. ist die Auflösung senkrecht zu den Schichten geringer als innerhalb der Schichtbilder. Um isotrope Voxeldaten zu rekonstruieren, werden zwei um ca. 90° gegeneinander verdrehte 3D-Bildstapel aufgenommen und mittels einer SRR-Methode zu einem hochaufgelösten 3D-Volumen rekonstruiert. Während der SR-Rekonstruktion werden

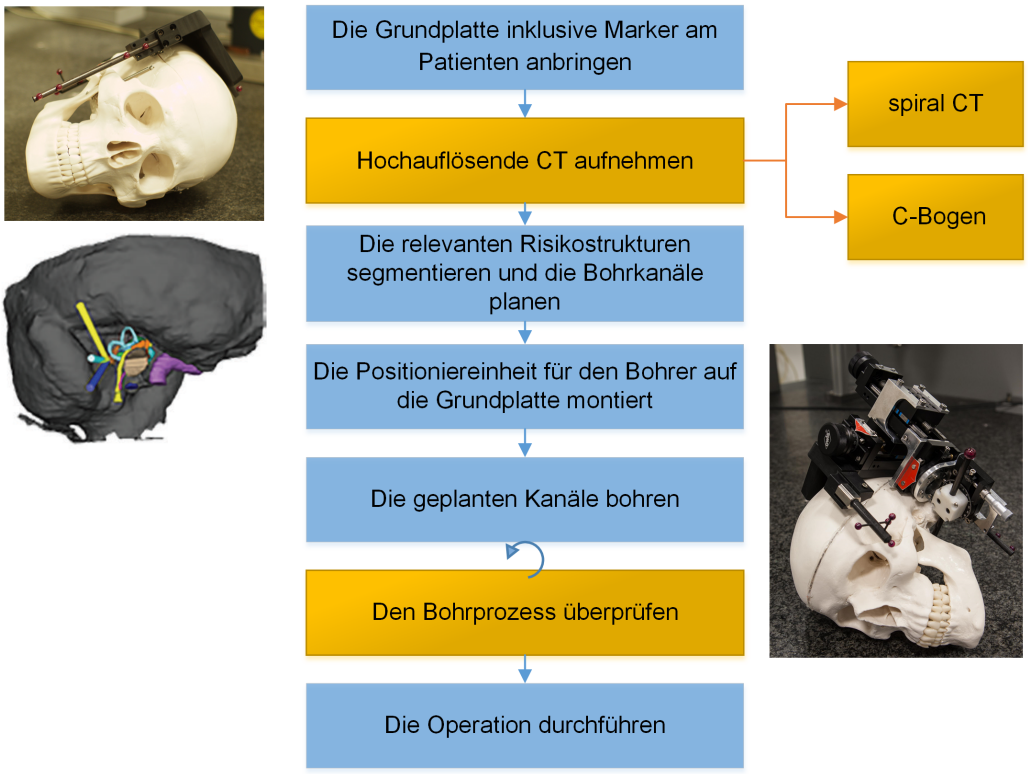


Figure 0.1.: Übersicht über die Prozessschritte für einen Multi-Port Eingriff an der Otobasis. Die zentralen Themen dieser Arbeit sind mit einem gelbem Hintergrund hervorgehoben.

die Input-Datensätze als niedrig aufgelöste Versionen des unbekannten hochaufgelösten Bildes betrachtet. Durch das Simulieren des Akquisitionsprozesses können niedrig aufgelöste Bilder generiert werden. Ziele der SRR-Methode ist dann ein Bild zu rekonstruieren, sodass der Unterschied zwischen originalem und simuliertem Bild minimiert wird. Aufgrund der niedrigen Anzahl der Input-Bilder (nur zwei Datensätze) sowie ihrer hohen Anisotropie (die Auflösung senkrecht zu den Schichten ist 3 bis 7 Mal schlechter als die Schichtauflösung) führen die klassischen Rekonstruktionsmethoden zu unscharfen Strukturen an den Stellen, wo Inputdaten fehlen. Weiterhin berücksichtigen klassische SRR-Ansätze die Unsicherheit der Input-Datensätze nicht. Daher wird in dieser Arbeit eine Methode entwickelt, die den Vertrauenswert der Inputdaten abschätzt [EW12] und diesen während der Rekonstruktion berücksichtigt [EW13]. Für eine verbesserte Rekonstruktion an der Stelle, wo keine Input-Daten vorhanden sind, wird eine neue Inpainting-Technik vorgeschlagen.

Die Rekonstruktionsmethode wird anhand synthetischer Daten als auch mit realen Scans evaluiert. Weiterhin wird der Einfluss der Bildverbesserung auf die Segmentierungsergebnisse bewertet.

Bildrekonstruktion aus Kegelstrahlprojektionen (CBCT)

Intraoperative 3D-Röntgenbildgebung ist zum Standard in der modernen Chirurgie geworden. Mithilfe von mobilen C-Bögen können 3D CT-ähnliche Bilder intraoperativ aufgenommen werden. Während der Rotation des C-Bogens um den Patienten werden 2D-Röntgenbilder akquiriert, die zur Rekonstruktion der Volumendaten verwendet werden. Die Bildqualität der CT Bilder ist jedoch aufgrund mehrerer Einschränkungen relativ schlecht.

Präzise geometrische Projektionsparameter (Position der Röntgenquelle, Detektor-Lage und -Orientierung) werden für jedes Röntgenbild benötigt, da sonst bei der Rekonstruktion Artefakte entstehen können. Aufgrund der mechanischen Stabilität des C-Bogens wird eine Offline-Kalibrierung üblicherweise durchgeführt, damit die Projektionsparameter ermittelt werden können. Es wird dabei angenommen, dass Abweichungen von der Idealgeometrie reproduzierbar sind; stochastische Abweichungen (z.B. Vibrationen bei der Rotation) werden dabei nicht berücksichtigt. In dieser Arbeit wird eine neue Methode zur geometrischen Online-Kalibrierung präsentiert, die zusätzlich die stochastisch-mechanischen Abweichungen kompensiert [ES15]. Die Qualität der Kalibrierung wird anhand von Simulations- und experimentellen Daten demonstriert.

Die Präsenz von Metallobjekten im durchstrahlten Bereich führen zu einer nichtlinearen Veränderung der Detektorantwort (Nutzsignal) und verursachen dadurch typischerweise Metallartefakte. Diese Artefakte können die Beurteilung von Bilddaten erschweren oder sogar unmöglich machen. Sowohl die standardmäßig verwendete gefilterte Rückprojektion (FBP) Methode als auch die iterativen Methoden, die einen linearen Rückprojektionskernel verwenden, erweisen sich als sehr empfindlich gegenüber der Totalabsorption der Röntgenstrahlung durch Metallobjekte. Um dieses Problem umzugehen, wird in dieser Arbeit eine nicht-lineare iterative Methode zur Reduktion der Metallartefakte entwickelt [ES14].

Intraoperative bildbasierte Instrumentenverifikation

Zur Kontrolle der aktuellen Instrumentenposition relativ zur Anatomie des Patienten werden mittels C-Bogen intraoperativ Bilddaten aufgenommen, um somit die Soll- und die Ist-Positionen während der Intervention zu vergleichen. Dafür muss die aktuelle Lage des Bohrers relativ zur Patientenanatomie bekannt sein. Um diese Aufgabe zu erledigen, werden zwei Durchleuchtungsbilder intraoperativ akquiriert. Hier werden nur zwei Bilder verwendet, um die Strahlenbelastung des Patienten gering zu halten. Nachdem der Bohrinstrument in beiden Bildern lokalisiert wird, kann anhand der entsprechenden geometrischen Projektionsparametern die 3D Lage des Bohrers berechnet werden.

Ein entscheidender Schritt dabei ist die Bestimmung der Lage des C-Bogens. Dafür wird eine Grundplatte mit Referenzstrukturen vor der CT-Aufnahme am Schädel des Patienten angebracht. Die Referenzstrukturen (Glass-Marker) sind sowohl im prä-operativen Volumen als auch in den intra-operativen Durchleuchtungsbildern sichtbar. Die Konfiguration der Marker ermöglicht die eindeutige Erkennung der Scanner-Lage relative zum Patienten.

In dieser Arbeit wird eine zweistufige Registrierung angewandt, um die Lage des C-bogens zu bestimmen: zunächst wird die Projektionsmatrix mit Hilfe der angebrachten Marker abgeschätzt. Die Marker-Mittelpunkten werden im proxi-operativen CT Volumen und in intra-operativen Durchleuchtungsbildern erfasst und basierend darauf wird die Projektionsmatrix abgeleitet. Anschließend erfolgt eine intensitätsbasierte Registrierung um die Projektionsparameter präziser zu ermitteln. Eine gute Initialisierung ist an der Stelle nötig, damit die Intensitätsbasierte Registrierung erfolgreich konvergiert. Die Hauptherausforderung dabei ist die voll-automatische Marker-Detektion und die präzise Ermittlung der entsprechenden Mittelpunkte im Durchleuchtungsbild, da aufgrund der Markerinhomogenität sowie möglicher Überlappung mit anderen Strukturen werden die Marker nicht als homogene Ellipsen abgebildet und können daher nicht genau lokalisiert werden. Um dieses Problem umzugehen, werden üblicherweise Metall-Marker eingesetzt, die als homogene dunkle Regionen abgebildet werden. Solche Marker verursachen aber

starke Metallartefakte in den rekonstruierten CT-Volumen und führen daher zur Verfälschung der Segmentierungsergebnisse. In dieser Arbeit wird eine robuste Methode vorgestellt, die Marker automatisch und präzise segmentiert auch im Fall einer Überlappung mit anderen Strukturen [EBS14].

Die Genauigkeit des Pose Estimation des C-Bogens wurde evaluiert, indem der Registrierungsfehler am Ziel (Target-Registration Error (TRE)) berechnet wird. Ein TRE von 0.167 $\pm$0.018 mm konnte erreicht werden.

Nachdem die Lage und Orientierung des Bohrinstruments berechnet sind, wird anschließend während des Vorschubs des Bohrinstruments ein visuelles Feedback bezüglich der Instrumentenspitze relativ zur Patientenanatomie angeboten.

Contents

1. Introduction

X-ray computed tomography (CT) has made a revolutionary impact on radiological diagnostics. It allows to acquire 3D images of the human body, which enables physicians to view internal organs. In several medical applications high-resolution (HR) images are required to allow accurate diagnose and perform precise treatment. Especially in the case of lateral skull base surgery which was the motivation of this work (DFG-supported project), precise 3D representations of the anatomy is of high importance. Thereby it is mandatory to avoid any damage of the critical anatomic structures lying in close proximity to the target region such as the cochlea and the facial nerve.

1.1. Clinical motivation

Minimally invasive procedures (MIPs) minimize the amount of tissue damage, lead to lower complication rates, and reduce postoperative pain as well as recovery time [SJA10]. Today, MIPs are well established in various surgical fields. However, the common clinical surgery of the lateral skull base remains largely traumatic: a large opening cavity is drilled by the surgeon in order to locate anatomical landmarks and preserve critical structures (see Figure 1.1). In addition, this process is time consuming and requires an appropriate skin incision. For instance, the common surgical technique of cochlear implantation is mainly based on mastoidectomy and posterior tympanotomy.

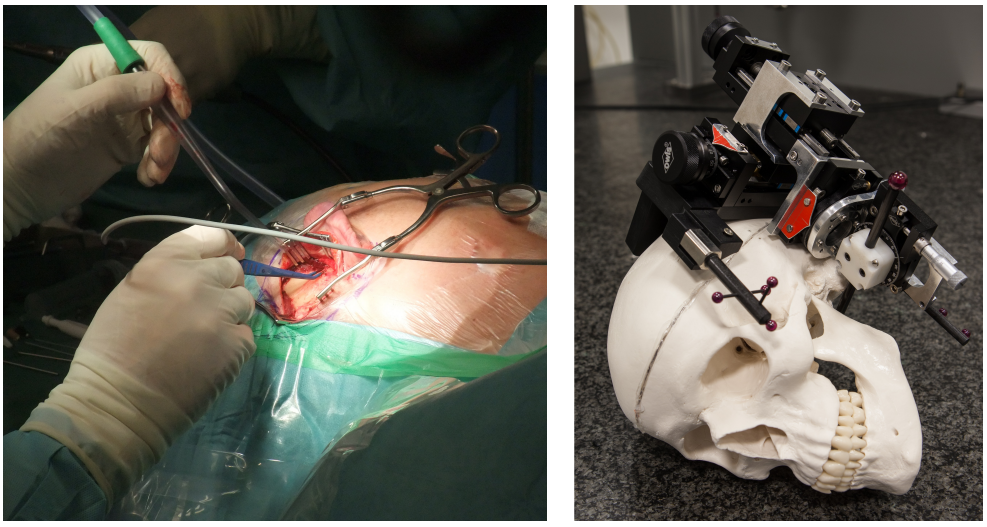


Figure 1.1.: Conventional largely traumatic skull base surgery is shown on the left. A prototype of a base plate used to perform minimally invasive procedure is depicted on the right.

In order to reduce invasiveness and substantial bone removal, several approaches for minimally invasive single port surgery have been attempted [GBG*14, HYSI13, LND*08, WBM*09]. A novel Multi-port minimally traumatic approach for lateral skull has been proposed by our DFG Research Group FOR 1585. The main idea is to drill tunnels which are slightly larger than the implant or the instrument to be used. Therefore, three canals are drilled: the first for the instrument, the second for the endoscope, and the third for material removal or for an additional instrument. Compared to single port procedures, the multi-port approach should enable advanced manipulation as well as direct visual feedback. In addition, larger space for manipulation can be created in the region of interest by exploiting the cavity resulting from the intersection of the drilled ports. These aspects are beneficial for typical interventions, such as tumor removal, biopsy, and cochlear implantation.

Clinical work-flow. In pre-operative stage, a base frame with reference structures (fiducial markers) is attached to the skull. These fiducials are employed for patient registration in the operation room. The patient's anatomy is then acquired with a computed tomography (CT) imaging modality, such as spiral CT. Based on the CT images, the critical anatomical structures are segmented and three drill tunnels are selected. A mechanical positioning device (to adjust the drill) is then mounted on the base frame, the initial pose of the drilling instrument is calculated and the drill process is started. X-rays images are then intra-operatively acquired with a C-arm CBCT system in order to control the drilling. After performing the planned tunnels, the treatment of the patient is accomplished and the wound is closed. This work-flow is illustrated in Figure 1.2.

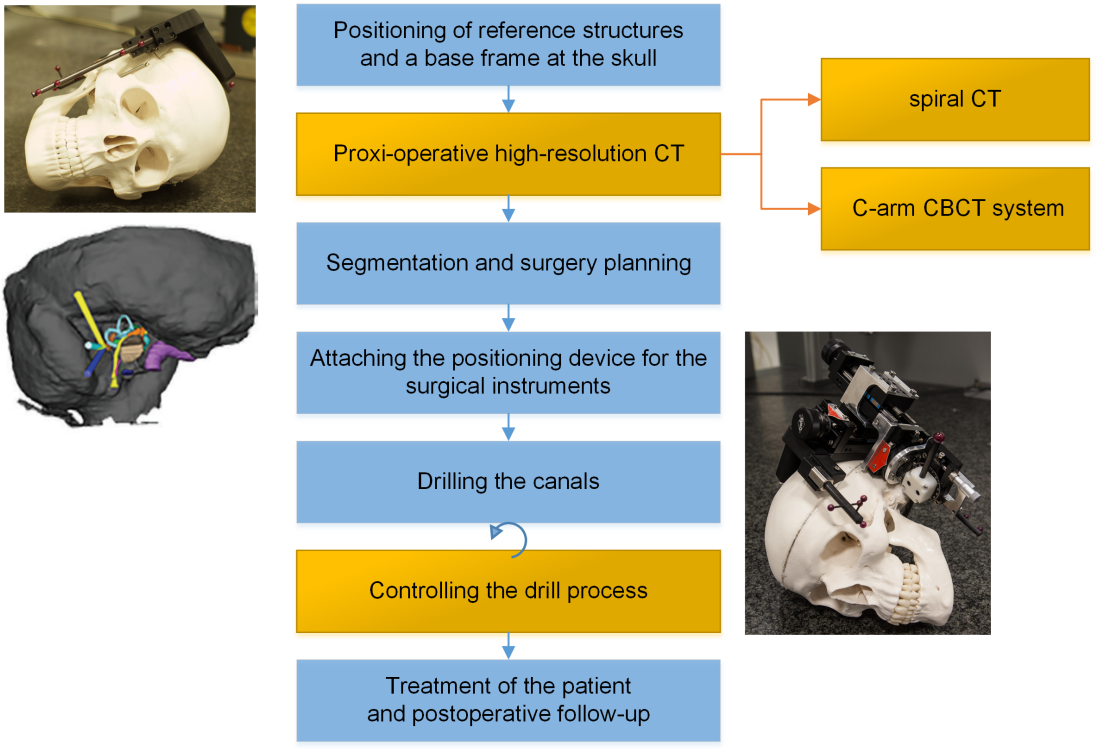


Figure 1.2.: Simplified process pipeline for image-guided minimally invasive surgery at the lateral skull base. The main topics addressed in this thesis are highlighted with yellow background.

1.2. Problems and contributions

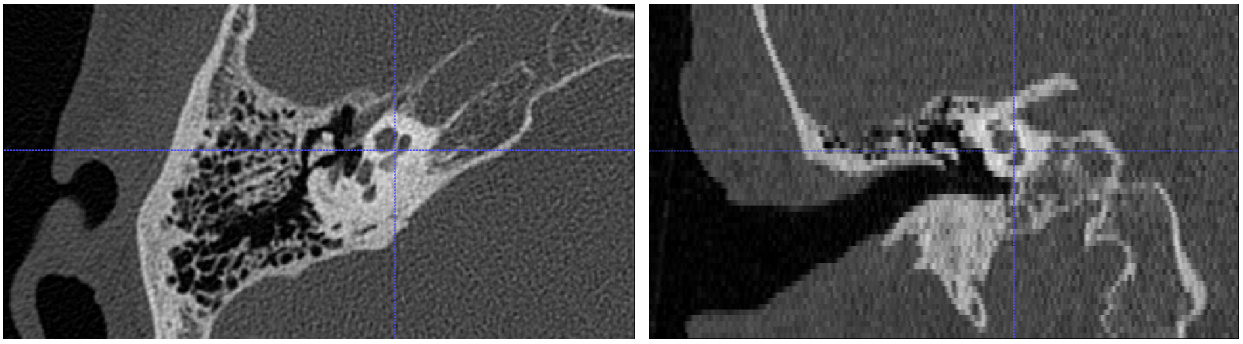
The imaging, segmentation, navigation and drilling process all have a major impact on the overall procedure accuracy. In this thesis we focus on two key factors: the high-resolution (HR) imaging of the region of interest by means of CT, and the image-guided navigation using a C-arm CBCT system.

1.2.1. High-resolution images

In order to acquire HR volumes with isotropic resolution, two techniques for two different imaging systems are investigated:

- indirect acquisition of HR 3D images using spiral CT scanners by applying super-resolution reconstruction methods
- direct reconstruction of HR volumes from 2D X-ray images gathered with a C-arm CBCT system

Super-Resolution Reconstruction (SRR). A direct acquisition of high-resolution data sets is often not feasible due to limitations in radiation dose and other technical constrains, such as acquisition time and limited detector cell size. Currently available spiral CT systems, especially single-detector row spiral CT, acquire data sets with large slice thickness, i.e. the resolution is high within the slices and low in the slice-selection direction (see Figure 1.3). Applying interpolation techniques can decrease the voxel spacing and visually smooth the result, but can not restore lost high-frequency details. In order to overcome these limitations, super-resolution reconstruction techniques can be applied.



(a) A representative axial slice of reconstructed 3D image having an isotropic in-plane spacing.

(b) A representative sagittal slice with an anisotropic spacing in the slice-selection direction.

Figure 1.3.: Illustration of resolution anisotropy in case of 3D images acquired with a spiral CT scanner. The slice-thickness is five times bigger than the in-plane spacing.

Super-resolution reconstruction refers to the set of methods aiming to recover high-frequency content, lost due to degradations in the imaging system, by combining several low-resolution (LR) images of the same scene, having slightly different image contents, into a high-resolution image. Given a model of the image acquisition process, SRR estimates the unknown HR image from the observed LR images.

We propose a novel super-resolution reconstruction algorithm aiming to reconstruct a HR image with isotropic resolution from two more or less orthogonal CT volumes which have anisotropic resolution. We fix the number of LR images to two in order to keep the patient radiation exposure in an acceptable range.

We propose an effective framework for fusing LR data sets under consideration of uncertainty of the input information. We present a novel enhanced directed interpolation-restoration technique to improve the initial guess of the reconstructed HR image. The proposed algorithms are shown to be fast and robust, and are therefore appropriate for clinical applications.

CBCT reconstruction. Spiral CT scanners are not appropriate for intra-operative imaging, and are therefore limited to pre- or post-operative image acquisition. In contrast, C-arm cone-beam computed tomography (CBCT) offers CT-like 3D imaging capabilities while being appropriate for interventional suites. Thereby, CBCT volumes are reconstructed based on X-ray images gathered during the rotation of the C-arm around the object under investigation. However, these systems are disturbed by higher mechanical inaccuracies compared to spiral CT due their open design. For instance, the gravity causes contraction and expansion of the C-arm during its rotation around the scanned object. Given that the source and detector do not follow an ideal scan trajectory, reconstruction artifacts are expected if such misalignments are not considered (see Figure 1.4). Under the assumption that the C-arm motion is reproducible, geometrical offline calibration is usually performed using an appropriate calibration phantom. However, by applying offline calibration techniques only systematic (reproducible) errors can be corrected. In order to compensate also stochastic misalignments occurring during the scan process, we propose an online calibration method which do not require any additional calibration phantom.

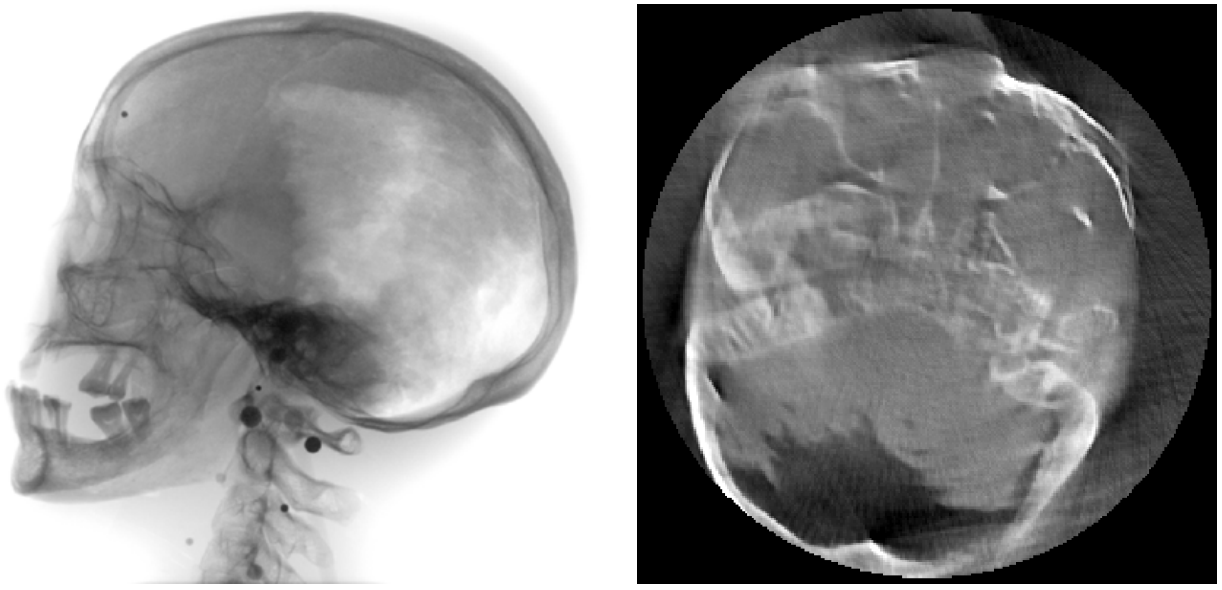


Figure 1.4.: Example showing motion artifacts: an X-ray image of a head phantom is shown on the left, and a representative slice of the corresponding 3D reconstruction disturbed by misalignment artifacts is on the right.

The success of CBCT is not only attributed to the technical advances but also to the improvements in the image reconstruction algorithms. Several approaches were developed to deal with different CT reconstruction artifacts which may disturb the interpretation of the acquired 3D images. Metal-related artifacts are reported to be one of the most challenging artifact types. They are caused by metal implants, and appear in form of shadows and streaks as depicted in Figure 1.5. We propose a novel iterative reconstruction algorithm aiming to eliminate such artifacts. In contrast to conventional metal-artifact reduction techniques,

we do not consider the metal projection regions as missing data. In addition, no prior knowledge about the form, position, or presence of metal implants is required. Despite the higher computational complexity of the proposed algorithm, it still be practicable for clinical use since it requires less than one minute for the whole reconstruction process.

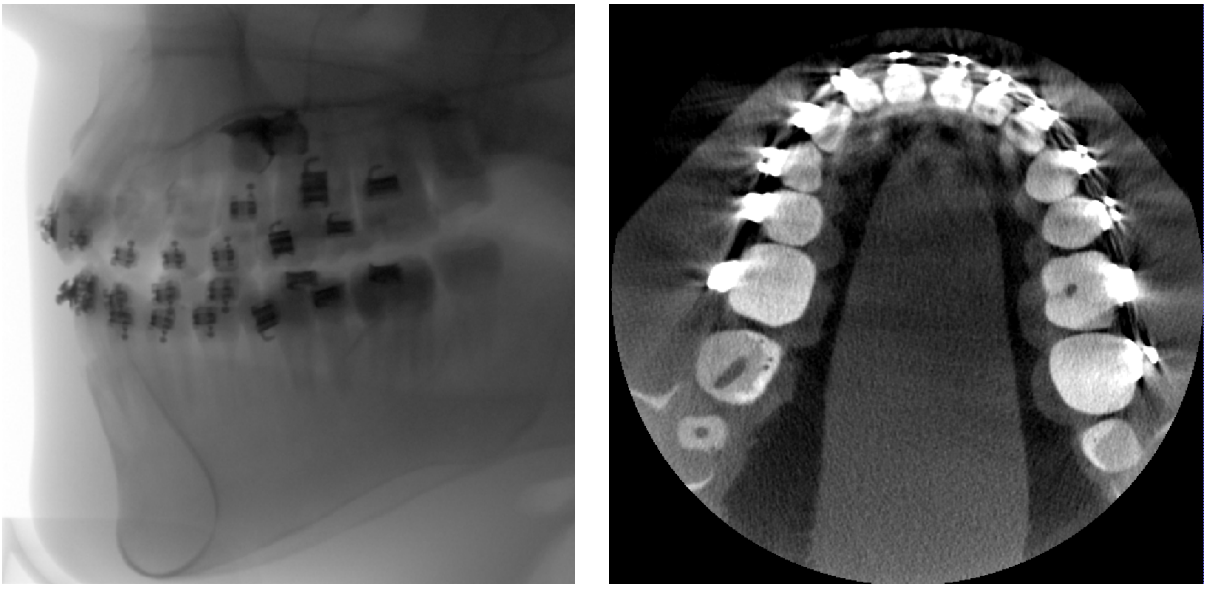


Figure 1.5.: Example illustrating metal-related artifacts: an X-ray image of a head phantom is shown on the left, and a representative slice of the corresponding 3D reconstruction exhibiting metal-related artifacts in form of streaks and shadows is on the right.

1.2.2. Navigation

An intra-operative control of the surgical instrument pose has to be performed in order to avoid possible damages of critical anatomic structures during the drill process. Tracking the drilling instrument can be achieved with an electromagnetic tracking system (EMT). However, the accuracy of EMT systems is relatively low, especially in the presence of metal objects. An image-based tracking can be done using a C-arm CBCT system. In this case it is mandatory to determine the pose of C-arm system relative to the surgical plan and the patient. This step is called registration and can be achieved using fiducial markers fixed to the skull.

It is well known that the target registration error depends on the fiducial localization error. Thus, it is very import to localize the fiducial markers in the X-ray images with sub-pixel accuracy. Several automated fiducial detection approaches have been presented. However, the reported registration uncertainty is too high to be applied for high-precision interventions. Moreover, usually titanium surgical screws or beads are used as markers. Such markers have high attenuation coefficients and cause therefore CT artifacts if conventional reconstruction algorithms, like FDK, are applied. Using markers made of other materials, such as glass, may be advantageous since they do not obscure anatomical structures in the superposition region in the 2D X-ray images. However, using glass markers may be challenging for the detection process due to their low contrast. We propose a novel segmentation approach aiming to define marker center points in X-ray images with sub-pixel accuracy even in case of low contrast or superposition with other structures.

After performing the registration and the segmentation of the drilling instrument in a fully automatic way, the deviation between the drilled canal and the appropriate planned path can be computed. For this task, only two X-ray images acquired with a C-arm CBCT system are needed to determine the spatial pose of the drill bit such that the risk of radiation exposure can be significantly reduced.

1.3. Outline of this thesis

This work is organized as follows: Chapter 2 deals with the super-resolution reconstruction. It reviews the state-of-the-art of SRR with a focus on its application in medical imaging. It describes new algorithms aiming to estimate the confidence of the input LR images, perform an initial guess of the unknown volume, and then iteratively reconstruct the HR image. A detailed evaluation of the proposed methods is given at the end of this chapter.

Chapter 3 starts with an overview of the physical principles underlying the CBCT and common CT reconstruction artifacts. Second, a state of the art of conventional CBCT reconstruction algorithms is presented. Third, new algorithms for online geometrical calibration, and for reconstruction of CBCT volumes with reduced CT artifacts are described. Finally, the proposed reconstruction algorithms are evaluated based on simulated and real data.

Chapter 4 deals with the image-based navigation. It covers approaches proposed for accurate registration of intra-operative fluoroscopic images to a pre-operative CT volume. After presenting methods proposed to localize the drilling instrument in X-ray images, and then calculate its spatial pose, the navigation process is described.

A conclusion of the thesis and recommends for future work are summarized in chapter 5.

Several parts of this thesis have been published in peer-reviewed conference proceedings and as journal publications. A list of these publications is given in Appendix A. The work has been influenced by several Diploma, Master's and Bachelor's theses that have been written under my supervision by students graduating at Technische Universität Darmstadt. These theses are listed in Appendix B. Finally, Appendix C gives a short curriculum vitae that covers my education and work experience.

2. Super-resolution reconstruction

2.1. Introduction

Since the discovery of the X-rays by Röntgen in 1895 [Rön96], it evolved within few years to a well established examination and treatment medical tool. By the 1960s, Hounsfield developed the computed tomography (CT) for the EMI Scanner [Hou80]: data is acquired from multiple X-ray transmissions through the object of interest and used to reconstruct a 2D image showing its inside. Technical advances leading to lower acquisition time, reduced dose usage and higher image quality allowed a wide spectrum of applications in medical imaging. For instance, the acquisition time decreased from several days in the first scanner to less than 0.2 seconds in modern systems. For more information on computed tomography, the reader is referred to [Kal11, Web90].

Most of image-based applications, such as diagnosis, planing and navigation, depend critically on the quality of the acquired CT scans. A main characteristic for high image quality is the resolution which is desired to be high and isotropic since higher resolution means more image details. Acquiring high-resolution (HR) images by improving the image acquisition device is not always possible due to several technical limitations. For example, a reduction in the detector cell size, i.e. a reduction of pixel spacing, may cause a degradation in the captured image quality since the signal power decreases proportional to the sensor's cell size reduction.

Spiral CT systems allow to continuously acquire entire anatomic regions in short time (2 to 30 seconds) and are therefore practical for clinical applications. However, the acquired data sets have usually an anisotropic resolution. As shown in Figure 2.1, it is to be distinguished between the in-plane resolution and the through-plane resolution (slice thickness). Usually, the in-plane resolution is 2 to 10 times higher than the through-plane resolution. Reducing the table shift step to achieve an isotropic resolution would result in higher radiation dose and longer scan time. Furthermore, by decreasing the spiral pitch, the generated volume is heavily blurred across all dimensions [Gre08]. Applying simple deblurring techniques can not achieve the desired high-resolution. Current spiral CT systems can scan the complete body trunk with an effective slice thickness of 0.1-0.2 mm. However, the resolution of reconstructed CT volumes is limited by the point spread function (PSF) of the spiral CT system, which has an equivalent bandwidth of 1-2 mm [Gre08]. Given all these limitations, a promising technique to acquire 3D images with isotropic resolution is the super-resolution reconstruction (SRR) .

Super-Resolution Reconstruction is the estimation of a high-resolution (HR) image from one or several low-resolution (LR) data sets. Due to incompleteness in the imaging process, image details may be lost in every single acquisition. Under the assumption that the acquired LR images are slightly different views of

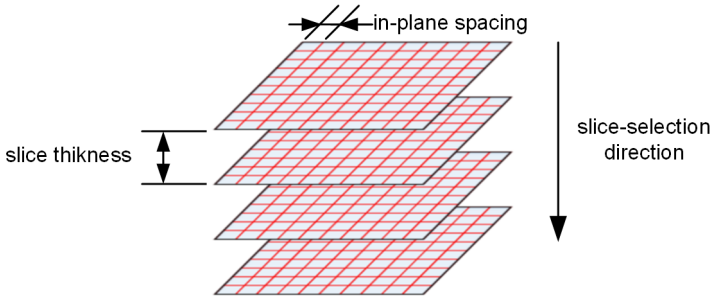


Figure 2.1.: Illustration of the anisotropic resolution of the 3D images acquired with spiral CT systems.

the same scene, the spatial resolution can be improved by combining information present in the input data sets.

Several SRR approaches have been proposed for photographic imaging. Applying these techniques to CT remains a challenging task due to different aspects. The goal in medical imaging is to reconstruct accurate images with less artifacts rather than visually pleasing imagery. Unlike photographic imaging, CT constrains illumination to prevent tissue damage and limits thereby the signal-to-noise ratio. Another key point is the number of the input images [BK02]. Due to the limited number of the anisotropic input images, lost image details have to be recovered from scattered anisotropic data. A crucial issue is therefore the reconstruction of data at positions where no sufficient input information are acquired.

This chapter presents a new approach to reconstruct a high-resolution CT image from only two highly-anisotropic low-resolution volumes. Therefore, confidence maps are introduced to estimate the data fidelity and reject outliers. A protocol for the acquisition of the LR images is proposed to prevent additional equipments which may be needed to control the shift between the observed data sets. Furthermore, a new method to enhance the initial HR estimate is presented. The main contributions of this chapter have been presented at IEEE CBMS 2013 [EW13], at the SPIE Medical Imaging 2012 [EW12], and have been published in [WBE11] (2011).

2.2. Background

2.2.1. Computed tomography

X-ray computed tomography (CT) aims at reconstructing a 3D presentation of the object under investigation by irradiating it from several angles. From each irradiation angle, a projection profile is acquired. This profile describes the attenuated X-ray intensity reaching the detector. Based on the projection profiles from all angles, it is possible to recover the attenuation coefficients of the scanned object.

2.2.2. Spiral CT

A spiral CT scanner consists of an X-ray tube and a detector continuously rotating around the scanned object while the table moves forward in the z-direction, which results in a helical scan trajectory as depicted in Figure 2.3. The initially developed single-row spiral CT was enhanced by installing multi-rows of detector cells helping to improve the longitudinal resolution. A further development is the use of z-sampling techniques. A comprehensive survey of developments in CT is given in [Kal11].

Figure 2.2.: A Spiral CT scanner [Sie].

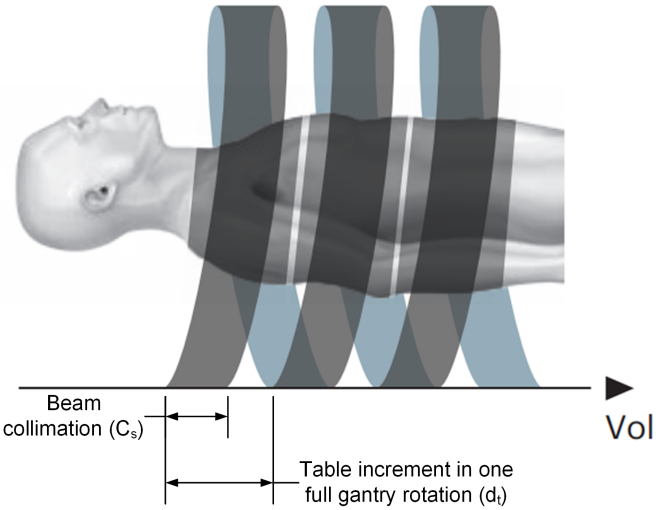


Figure 2.3.: An illustration of helical scan trajectory in spiral CTs with a pitch factor >1.0 [Sie12].

The helical pitch (p) is one of the principle terminologies in spiral CT. It is defined as the table distance (d_t) traveled in one full gantry rotation divided by the beam collimation (C_s) as illustrated in Figure 2.3:

$$p = \frac{d_t}{C_s} = \frac{d_t}{N_s \cdot T_s} \tag{2.1}$$

where N_s is the number of detector rows and T_S is the nominal slice thickness. In case of single-slice spiral CT, the slice thickness T_S and X-ray beam width C_s are the same, unlike multi-slice CT whose beam collimation is equal to the total thickness of all simultaneously acquired slices. For instance, for a 64-slice CT scanner, i.e. $N_s = 64$, with a nominal slice thickness $T_s = 0.5$ mm and a table feed $d_t = 32$ mm per rotation, a pitch $p = 1.0$ results.

The pitch factor affects the image quality and the patient dose. A larger table increment d_t results in faster acquisitions, but leads to lower resolution. If pitch factor is equal to one, the X-ray beams are contiguous for adjacent rotations. Increasing the pitch ($p > 1.0$) results in decreased patient dose, but also in gaps between the X-ray beams (see Figure 2.3), which degrades the image quality. In contrast, decreasing the pitch ($p < 1.0$) may enhance the image quality, but on cost of patient dose. Achieving approximate isotropic resolution could be obtained with a very thin nominal slice thickness ($T_s < 0.5$mm) at a pitch $p = 1$. In clinical applications, higher nominal slice thickness ($T_s > 1mm$ [LF15]) is used due to large scanning ranges and patient dose limitations, which results in a considerable difference between the in-plane and the out-of-plane (longitudinal) resolution.

2.2.3. Ill-posedness and regularization

Given a set of observed images and an appropriate model of the imaging process, the super-resolution reconstruction tries to estimate the corresponding high-resolution (HR) image. Thus, the reconstruction process is defined as an inverse problem. In the following some aspects related to the nature of this investigated inverse problem are described.

Ill-posedness. The SRR problem is an ill-posed one since its solution doesn't satisfy at least one of the Hadamard requirements [Had23]. Due to the characteristics of the LR images and the imaging process, the solution may not

- exist for any data: the observed low-resolution images may be inconsistent with any high-resolution image if they are corrupted by noise. Consequently, the problem is non-invertible and no HR image can be estimated from the observed images.
- be unique: a non-trivial solution space which is consistent with the observed LR images may exist if the imaging process have a surjective operator. Thus, the solution of this problem is expected to be non-unique. Furthermore, the provided independent constrains are typically insufficient to recover a unique solution.
- depend continuously on the data: The restoration process may be sensitive to data inaccuracies. For example, if the imaging system has a spectral response decreasing asymptotically toward zero with increasing frequency, a small perturbation at high frequency leads to a large spurious signal in the reconstructed image.

The ill-posedness problem can be solved through regularization. Therefore, inclusion of prior knowledge about the solution characteristics may be very significant since it provides valuable constrains for the solution space.

In practice, the observed data and the solution are discretized. In case of ill-posed problems this discretization leads to ill-conditioned numerical problems. Thus, regularization methods are required to obtain acceptable solutions.

Regularization. In order to compensate for the information loss characterizing ill-posed problems, additional information is required. This later is typically provided in form of a-priori (or prior) information which cannot be derived from the observations or the acquisition process. The prior information aims to reduce or constrain the space solution by setting characteristics for the desired solution such as smoothness, positivity, total energy, etc. This results in formulating a new well-posed problem closely related to the original ill-posed one. The unique solution of the new well-posed problem has to be meaningful with respect to the original problem [Han98]. Thus, a-priori constrains reflecting the characteristics of the solution must be accurately defined. For instance, Hong et al. [HKK97] proposed the use of Tikhanov regularization [TA77] for SRR, which imposes smoothness constrains on the feasible solution space.

2.3. Related work

Super-Resolution Reconstruction is the estimation of high-resolution image typically from several low-resolution images. It has been used in many applications where multiple images of the same scene can be acquired such as in satellite and astronomical imaging, multi-sensor image fusion for surveillance applications, etc. Some approaches were proposed to extend the super-resolution framework to medical image applications like diagnosis and planing. This section reviews the state-of-the-art of the SRR and its application in the medical imaging.

Super-resolution reconstruction techniques are broadly categorized into two groups: single- and multi-input reconstruction methods. The first category is known as Single-Input Single-Output (SISO) SRR. It covers approaches which reconstruct a single output image from a single input data set. The performance

of these methods is limited by the information present in the observed single input. In contrast, additional observation constraints are provided in case of multi-inputs SRR, which may improve the reconstruction performance.

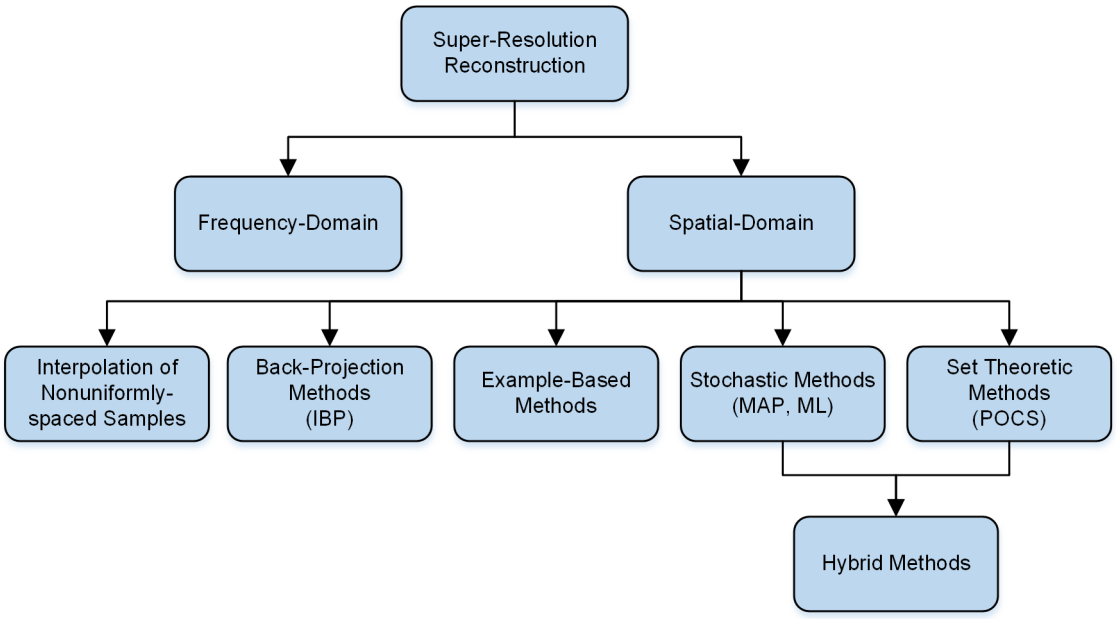


Figure 2.4.: Classification of the super-resolution reconstruction methods.

A second classification of SRR techniques can be performed based on the reconstruction domain, namely into the frequency-domain and the spatial-domain approaches (Figure 2.4). The frequency-domain methods have the advantage of being simple and computationally attractive. However, they do not allow the inclusion of prior knowledge or/and noise models. In contrast, spatial-domain methods offers the possibility to include prior knowledge. The most significant approaches among the spatial domain methods are the Iterative Back-Projection methods (IBP) [PKS87, KPB88, IP90, IP91], the Maximum A-posteriori Probability (MAP) estimation method [RR96, HBA97, TK94, EF97], and the Projection Onto Convex Sets (POCS) method [SO89, TOS92, DW82].

In the MAP approach, the SRR is considered as a statistical inference problem. Prior knowledge is used in the form of a-priori probability density on the solution. In the POCS method, the solution space is defined as the intersection of a collection of convex constrain sets. Hybrid methods combining MAP and POCS have been also proposed. Table 2.1 summarizes the comparison between the two main classes of the SRR techniques. In general, the spatial-domain reconstruction approaches may provide better results but at cost of theoretical and computational complexity.

2.3.1. Frequency-domain approaches

The frequency-domain SRR methods employ especially the shifting property of the Fourier transformation and the sampling theorems. This technique has the drawback of precluding the use of general observation and motion models. As an example, a spatial variation in the spatial-domain has no duality in the Fourier-domain and can therefore not be considered.

	Frequency-Domain	Spatial-Domain
Motion models	global translation	almost unlimited
Observation model	limited	flexible
A-priori constraints	limited	almost unlimited
Regularization	limited	excellent
Simplicity	✓✓	✗
Extensibility	✗	✓✓
Computational cost	✓✓	✗
Performance	✓	✓✓

Table 2.1.: Comparison of the frequency-domain and spatial-domain super-resolution reconstruction techniques.

The earliest work was proposed by Tsai and Huang in 1984 to enhance the resolution of satellite images [TH84]. The motion blurring effect and the observation noise were ignored. The inter-frame motion were limited to translational case. The authors used the shift and aliasing properties to formulate an equation system which relates the observed LR images (in form of aliased discrete Fourier transform coefficients) to the unknown HR image (in form of continuous Fourier transform samples). By solving this system and then using the inverse DFT, the HR image can be recovered. Therefore, the inter-frame motion has to be estimated in order to formulate the system of equations. The authors proposed a simultaneous multi-frame image registration method which should deliver reliable registration results for under-sampled images. This algorithm requires a sufficiently large number of observed images and is limited to the estimation of a global translation motion. Extensions to this method have been proposed by Tekalp et al. [TOS92] to include a linear, shift invariant Point Spread Function (PSF) as well as observation noise. The problem is formulated analog to the work of Tsai and Huang [TH84], but the solution is determined in sense of the least squares due to presence of noise. The issue of ill-posedness caused by the presence of zeros in the spectral representation of the blur kernel was addressed by Kim and Su [KS93]. The authors propose to apply a Thikanov regularization [TA77] based on squared error. This approach leads to reconstruct smoothed images.

In conclusion, frequency-domain methods have the advantages of simplicity and low computational complexity, but they have these significant drawbacks:

- Due to fundamental limitations, only global translational motion models can be employed.
- Difficulty in including spatially varying degradation models.
- Limitations in including spatially varying a-priori constrains.

These shortcomings are overcame in the spatial-domain multi-frame SRR [TM11], but usually at the cost of computational performance.

2.3.2. Spatial-domain approaches

The spatial-domain SRR methods offer more flexibility in defining the observation models and including prior information. Concerning the observation models, it is possible to model:

- global or non-global arbitrary motion,
- degradation, such as optical system degradation,
- non-ideal sampling,
- and motion blurring.

2.3.2.1. Interpolation of non-uniformly-spaced samples

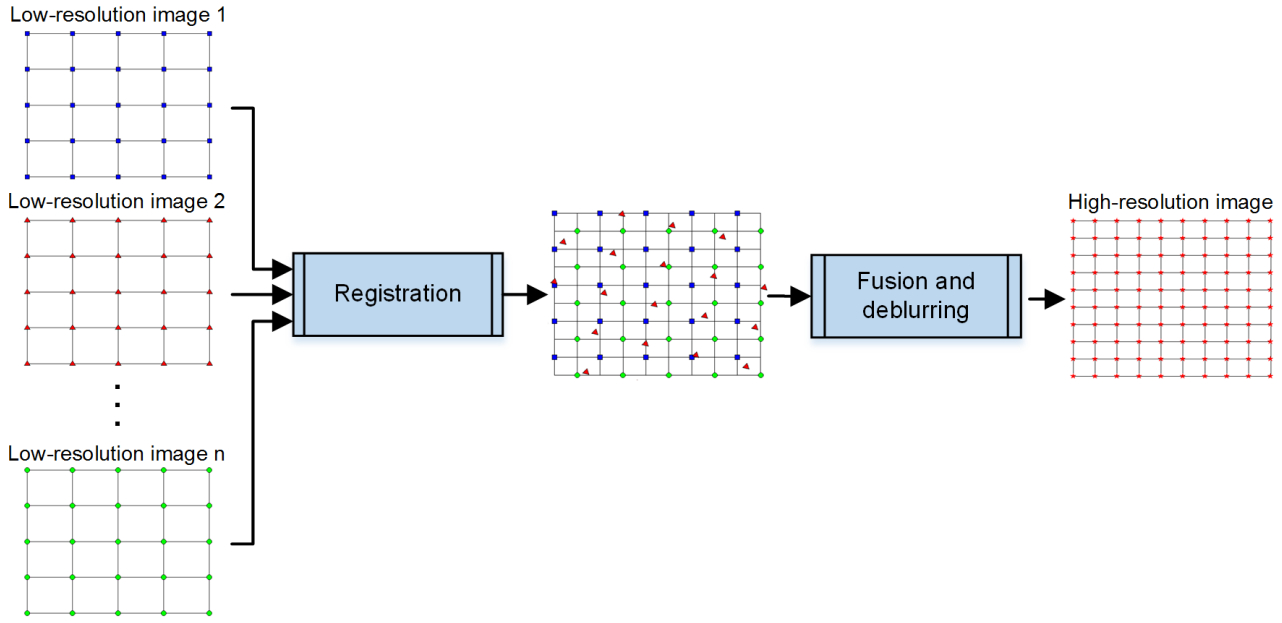


Figure 2.5.: Illustration of the interpolation-based super-resolution reconstruction.

Main processing steps of the interpolation-based SRR methods are illustrated in Figure 2.5. First, a registration is performed to align observed images to a fixed reference frame. A composite image of all non-uniformly-spaced samples is then generated. The non-uniformity is due the arbitrary shift between the low-resolution images. To restore a uniformly-spaced HR image, the non-uniformly-spaced samples are interpolated and resampled on a regular high-resolution lattice. The interpolation-based technique is simple and intuitive, but have the shortcoming of simplifying the imaging model by ignoring blurring and noise. Thus, it tends to not preserve edges and produce blurring artifacts.

All SRR methods perform at some stage non-uniform interpolation, sometimes in form of projection. Common algorithms for image interpolation are nearest neighbor, bilinear, and B-spline interpolation [UAE91]. Given a more realistic imaging model, better reconstruction results can be achieved. Ur and Gross [UG92] proposed a method which performs non-uniform interpolation followed by deblurring to reduce the blur artifacts. A wavelet-based algorithm was presented by Nguyen and Milanfar [NM00]. It is worthwhile to note that the image interpolation alone can not reconstruct the lost high-frequency components.

2.3.2.2. Iterative Back-Projection methods

Some SRR methods are based on computed tomography reconstruction algorithms, namely the Iterative Back-Projection (IBP). These approaches comprises basically two steps: simulate the observed images and then correct the estimated HR image. Assuming that the observation model as well as an estimate of the HR image are available, it is possible to simulate LR images. By comparing the simulated LR images with the observed images, a residual error is computed. If the estimated HR image as well as the imaging model are accurate, the residual error tends to be small. By updating the SR estimate in order to minimize the residual error, it is expected to enhance the reconstructed image.

Given a linear observation model H, a low-resolution image Y can be derived from a high-resolution image X following Equation 2.2.

$$Y = HX \tag{2.2}$$

where X and Y are the HR and the LR images, respectively.

By applying Equation 2.2, a low-resolution image $\widehat{Y}$ can be simulated from an estimate $\widehat{X}$ of the unknown high-resolution image X as

$$\widehat{Y} = H\widehat{X} \tag{2.3}$$

The estimate of the HR scene is updated by back-projecting the difference between the simulated LR images $\widehat{Y}$ and the observed images Y. The back-projection operator H^{BP} is in general an approximate of the inverse of the operator H. The SR image $\widehat{X}$ at iteration k is computed as

$$\begin{aligned} \widehat{X}^k &= \widehat{X}^{k-1} + H^{BP}(Y - \widehat{Y}^{k-1}) \\ &= \widehat{X}^{k-1} + H^{BP}(Y - H\widehat{X}^{k-1}) \end{aligned} \tag{2.4}$$

Main difference to the IBP in computed tomography is the choice of the projection and back-projection operators. Peleg et al. [PKS87, KPB88] presented an IBP approach for super-resolution reconstruction to restore unchanging 2D scene. The authors used a global translation and rotation model to estimate the motion parameters. The blurring is modeled as a spatially invariant PSF. The proposed back-projection operator is simplistic and makes very limited use of prior knowledge about the imaging system. This method was extended to improve the back-projection process [IP90, IP91]: analog to CT, the back-projection kernel scales the correction to the HR estimate depending on the contribution of a given HR pixel on the corresponding LR pixels. Thereby, no observation noise was modeled under the assumption that small amounts of noise would not drastically affect the solution, but this assumption is often not fulfilled. IBP methods provide a mechanism to constrain the reconstructed SR volume to the observed images, but this mechanism is only as good as the observation model.

2.3.2.3. Stochastic methods

These methods use stochastic models to describe the imaging process: the unknown HR image and the corresponding observations are not treated as deterministic systems but rather as stochastic models. The regularization of the ill-posed SRR problem is performed based on statistical estimation theory like Bayesian techniques. The Bayesian method, also known as maximum a-posteriori (MAP) estimation, can include prior constrains in form of a-priori probability density functions on the unknowns.

The observation model relating the unknown HR scene X to the LR image Y is defined as

$$Y = HX + \upsilon \tag{2.5}$$

where υ is an additive noise. The MAP method estimates $\widehat{X}$ approximating X by maximizing the a-posteriori probability $\mathcal{P}\{X|Y\}$:

$$\widehat{X} = \arg\max_X \left[\mathcal{P}\{X|Y\}\right] \tag{2.6}$$

The Equation 2.6 is extended using the Bayes' rule to

$$\widehat{X} = \arg\max_{X} \left[\frac{\mathcal{P}\{Y|X\}\,\mathcal{P}\{X\}}{\mathcal{P}\{Y\}} \right]. \tag{2.7}$$

The joint probability is independent of X in the denominator. Thus, the MAP can be simplified to

$$\widehat{X} = \arg\max_{X} \left[\mathcal{P}\{Y|X\}\,\mathcal{P}\{X\} \right] \tag{2.8}$$

By applying the –monotonic increasing– logarithm function, the Equation 2.8 is reformulated to maximize the sum of two terms: the log-likelihood function ($log\mathcal{P}\{Y|X\}$) and the logarithm of the a-priori density ($\mathcal{P}\{X\}$), as

$$\widehat{X} = \arg\max_{X} \left[\log \mathcal{P}\{Y|X\} + \log \mathcal{P}\{X\} \right] \tag{2.9}$$

The log-likelihood function describes the constrains relating the LR data to the unknown HR image.

According to the Equation 2.5, it is evident that the likelihood function is determined by the noise probability density $\mathcal{P}_v$ as

$$\mathcal{P}\{Y|X\} = \mathcal{P}_v(Y - HX) \tag{2.10}$$

In order to ensure the existence and uniqueness of the solution $\widehat{X}$, the noise is often assumed to be Gaussian and the prior model to be a convex function of X. In case of a zero-mean Gaussian noise model with variance σ^2, the Equation 2.10 reduces to

$$\mathcal{P}\{Y|X\} = \frac{1}{(2\pi)^{M/2}\sigma^M} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{n=1}^{N} \| Y_n - H_n X \|^2 \right\} \tag{2.11}$$

where M is the number of the HR image pixels, and N is the number of the LR images.

Prior model. The prior probability $\mathcal{P}\{X\}$ of the data model should reflect the characteristics of the imaging process. Usually, the prior distribution is defined as a Markov Random Filed (MRF) that is characterized by its Gibbs distribution. Giving a set of cliques $\mathcal{C}$ for the MRF and a potential function φ_c, an energy function Ω can be defined as

$$\Omega(X) = \exp\left\{ -\sum_{c\in\mathcal{C}} \varphi_c(X) \right\} \tag{2.12}$$

Typical prior models are

- white noise prior: $\Omega = \|X\|_2^2$ (L_2-norm)
- smoothness prior: $\Omega = \left\|\nabla^2 X\right\|_2^2$ (L_2-norm of Laplacian)
- impulsive data prior: $\Omega = \|X\|_1$ (L_1-norm)
- total variation prior: $\Omega = |\nabla X|$ (magnitude of gradient)

Based on the Gibbs prior model, $\widehat{X}$ is estimated by minimizing the following cost function

$$\widehat{X} = \arg\min_{X} \left\{ \| Y - HX \|^2 + \lambda\Omega(X) \right\} \tag{2.13}$$

The regularization parameter λ weights the influence of likelihood term and the Gibbs prior term. If λ is set to zero, so that no prior term is considered, the MAP estimation reduces to Maximum Likelihood (ML) estimation. In contrast, an overweight of the prior term will reduce the fidelity to the observed LR images.

Schultz and Stevenson proposed to use Huber Markov Random Field (HMRF) prior model which is a discontinuity-preserving image model [RR96]. It imposes smoothness constrains on the reconstruction while preserving visually significant edge features. Hardie et al. [HBA97] presented a similar approach with the extension of estimating both the unknown HR image and the motion parameters $\hat{s}$ simultaneously. The optimization is formulated as

$$< \widehat{X}, \hat{s} >= \arg\max_{X,s} \left[\log \mathcal{P}\left\{X, s | Y\right\}\right] \tag{2.14}$$

In this case, the optimization over $\widehat{X}$ and $\hat{s}$ is done alternately. The motion is no longer directly estimated from the LR images but rather by minimizing the error between the simulated LR images and the observed data.

A special case of MAP estimation is the Maximum Likelihood (ML) estimation. Some approaches based on ML formulation have been proposed in [TK94, EF97]. These methods can not include prior knowledge and are therefore poorly suited to solve the ill-posed inverse SRR problem.

2.3.2.4. Set Theoretic methods

The method of Projection Onto Convex Sets (POCS) is very intuitive and simple to implement. It allows the inclusion of prior information by defining convex constraint sets which impose restrictions to the solution space. The main idea is to reconstruct a high-resolution image which simultaneously satisfies the totality of pre-defined constrain sets. The first application of POCS based SRR was presented by Stark and Oskoui [SO89], although the noise was not considered in the image formulation model. An extension to include sensor noise has been proposed by Tekalp et al. [TOS92].

The POCS is an iterative procedure aiming at finding a point (in the vector space) which satisfies all pre-defined constraints [DW82]. The constrains are formulated as convex sets encapsulating characteristics related to smoothness, bounded energy, positivity, etc. The solution space is the intersection of all of these constraint sets. Given p constraint sets S_q and appropriate projection operators P_q, the estimate of X at iteration k is updated by applying a sequence of projections as

$$X^k = P_1 P_2 \cdots P_p X^{k-1} \tag{2.15}$$

A commonly used constrain in the SRR is the consistency with the observed data, defined as

$$S_C = \left\{\widehat{X} : |Y_n(i) - (HX)(i) \leq \tau_n|\right\} \tag{2.16}$$

where i is the voxel position in the LR image grid and τ_n a threshold which can be freely set based on the noise power, uncertainty of motion parameters, etc.
Another common constraint is the bounded energy:

$$S_E = \left\{\widehat{X} : \|X\|_2^2 \leq E\right\} \tag{2.17}$$

A useful constraint to limit the solution to lie within some distance ϵ_R from a learned reference image X_{ref} is

$$S_R = \left\{ \widehat{X} : \|X - X_{ref}\| \leq \epsilon_R \right\} \tag{2.18}$$

Other constraints can be defined in similar manner.

The main challenge with the POCS technique is the determination of appropriate projection functions that ensure the convergence against the intersection of the defined constrain sets. In addition, it requires a large number of iterations to achieve convergence and its computational cost is very high [TM11].

2.3.2.5. Example-based Approaches

In case of single frame SRR, only one LR image is available. Example-based methods have been proposed to overcome the problem of missing observations. These methods try to learn the correlation between LR and HR details from training data sets. Freeman et al. [FJP02] started from a collection of high-resolution images and degraded each of them in order to generate a training set. Based on the created training data, correspondences between HR and LR patches are stored. Glasner et al. [GBI09] proposed an SRR approach for natural images. Their method employs recurrence of patches within the same scale as well as within different scales of the same image to recover the HR patches. The main drawback of the example-based technique is that it limits the solution to the learned patches, which may results in artifacts in case of new unlearned data.

2.3.2.6. Hybrid methods

The general idea of these methods is to combine advantages of stochastic and POCS methods to incorporate prior knowledge while constraining the solution space. Schultz and Stevenson [RR96] proposed a constrained MAP-SRR which is an alternation of POCS and MAP iterations. Elad and Feuer [EF97] introduced a hybrid ML-POCS method. The ML is used to formulate the SRR as a statistical inference problem, while the POCS effects the regularization.

2.3.3. Medical imaging applications of Super-Resolution Reconstruction

Magnetic Resonance Imaging (MRI). The super-resolution framework has to be customized based on the application proprieties. Initial attempts to adapt SRR techniques to medical imaging dealt with the magnetic resonance imaging [Fia01, PY01]. Peeters et al. [PKN*04] proposed to increase the resolution in functional MRI in slice direction by applying SRR to acquired slice shifted images. Rahman and Wesarg [RW10] applied MAP estimation to reconstruct images with isotropic resolution by combining multiple orthogonal cardiac MRI data sets.

A similar method to Galsner et al. [GBI09] has been proposed by Rousseau et al. to synthetic MRI brain [RKS10a]. In this method, the high-resolution images are reconstructed based on atlases of images of different subjects by applying a local fusion of similar patches. Shi et al. [SCL*13] proposed an approach to reconstruct cardiac images using multi-atlas. Their approach attempts to find patch correspondences between short axis MR images and a set of atlases in order to recover the HR patches. Shortcoming of all these example-based methods is the necessity to training data sets.

We have proposed an example-based method where only a single image is required [KHW13]. Therefore, we employed the high resolution of the acquired axial slices to improve the low resolution in perpendicular direction. In addition, the recurrence of small patches is performed across different scales of the same image. A main drawback of this example-based method is the restriction to the learned patches. Thus, artifacts may be generated in case of new unlearned details.

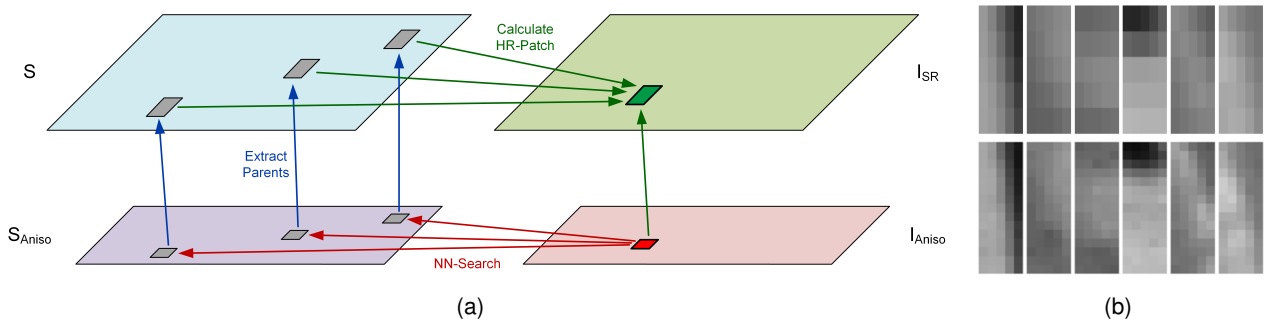


Figure 2.6.: Illustration of anisotropy correction by employing patch similarity (a). The images on the left (b) are examples of corresponding high-resolution and low-resolution patches [KHW13].

Positron Emission Tomography (PET). Several super-resolution techniques have been introduced to the positron emission tomography (PET). Kennedy et al. [KIF*06] proposed to acquire the LR data sets by shifting and rotating the detector and then apply the SRR method of Irani and Peleg [IP91]. Jeong et al. [JKKR13] presented a sinogram-based algorithm, which aims to reconstruct a HR sinogram from the measured LR sinograms. The authors used Monte Carlo simulations to estimate space-variant blur matrices and use them to restore the HR sinograms.

Mammography. Robinson et al. [MSC*11] proposed a wavelet-based approach to improve the resolution for digital X-ray mammography. The authors introduced a second framework to improve the azimuthal resolution of Optical Coherence Tomography (SDOCT) images. Their method consists of acquiring several repeated fast sparse 3D scans, remove frames affected by motion artifacts and then fuse the artifact-free scans.

Computed Tomography (CT). Only few a works on the application of SRR in computed tomography were presented so far. In the approaches of Bodensteiner et al. [BDS09] as well as Bernhardt et al. [BLD06], the reconstruction was limited to 2D case. They proposed to improve the resolution of the acquired 2D cone-beam CT (CBCT) projections by acquiring several images of the scanned object and then apply the SRR principles. Some SRR approaches have been proposed for lung 4D-CT [YGP*12, WXZ14]. Liu et al. [LYM*10] proposed a method based on applying SRR technique in sinogram domain. Main drawback of this method is that it requires the projection raw data and depends directly on the hardware in use. In addition, all of these approaches performed experiments based only on simulated data.

In summary, the mentioned super-resolution techniques have several limitations:

- The uncertainty in the observations is limited to the observation noise. Missing data or image artifacts are not considered
- Spatially varying PSF is limited to model the degradation caused by sampling with a low resolution sensor or relative motion between the scene and the sensor, but not the degradation caused by the reconstruction process which is applied to reconstruct the LR images
- Most of the proposed SRR methods simplify the image formation by assuming linear models

2.4. Concept

It was shown in several studies that iterative back-projection SRR results in sharp reconstructions with higher resolution compared to other SRR approaches since it does not smooth the image for regularization purposes [Ple13, IP91]. This approach is very stable and converges rapidly, which makes it suitable for clinical applications. However, it is sensitive to noise and outliers. This thesis contributes a new SRR approach which is based on the IBP reconstruction concept proposed by Irani and Peleg [IP91].

The proposed approach consists of the following steps:

1. Upsample the input anisotropic LR images in order to obtain isotropic voxels
2. Align the upsampled images to each other by employing a rigid registration with one volume as reference (see Section 2.4.3)
3. Compute uncertainty information for each input LR image as described in Section 2.4.4
4. Extract voxels with low uncertainty (smaller then a given threshold τ_ϵ) from the input images and merge them to one volume having the desired voxel spacing. This step creates a volume which consists of reference voxels with high confidence and free gaps in-between (sets of voxels with high uncertainty).
5. Fill the gaps by applying an oriented interpolation to create an initial estimate of the HR image (see Section 2.4.5)
6. Perform an iterative reconstruction to enhance the initial guess as detailed in Section 2.4.6

These steps can be summarized in three main stages: registration (1,2), initial estimation (3,4,5) and iterative restoration (6). An overview of this algorithm is shown in Figure 2.7.

Three main limitations with direct influence on the reconstruction quality are considered in this thesis:

- necessity to relatively high number of input LR images in case of high anisotropy (>3)
- limitation of the uncertainty in the observations to the noise in the imaging process: missing data and image artifacts are not considered
- limitation in the PSF to model the degradation caused by the sampling with a LR sensor and/or the relative motion between the scene and the sensor. Thereby, the degradation caused by the reconstruction process itself is not considered

The following assumptions are met:

- no motion blur caused by patient motion during the scan
- rigid registration is sufficient to align the acquired images. An extension to deformable structure can be achieved by using deformable registration method in the registration step. This extension is straightforward, however not required for the skull-base surgery and is therefore not implemented in this work

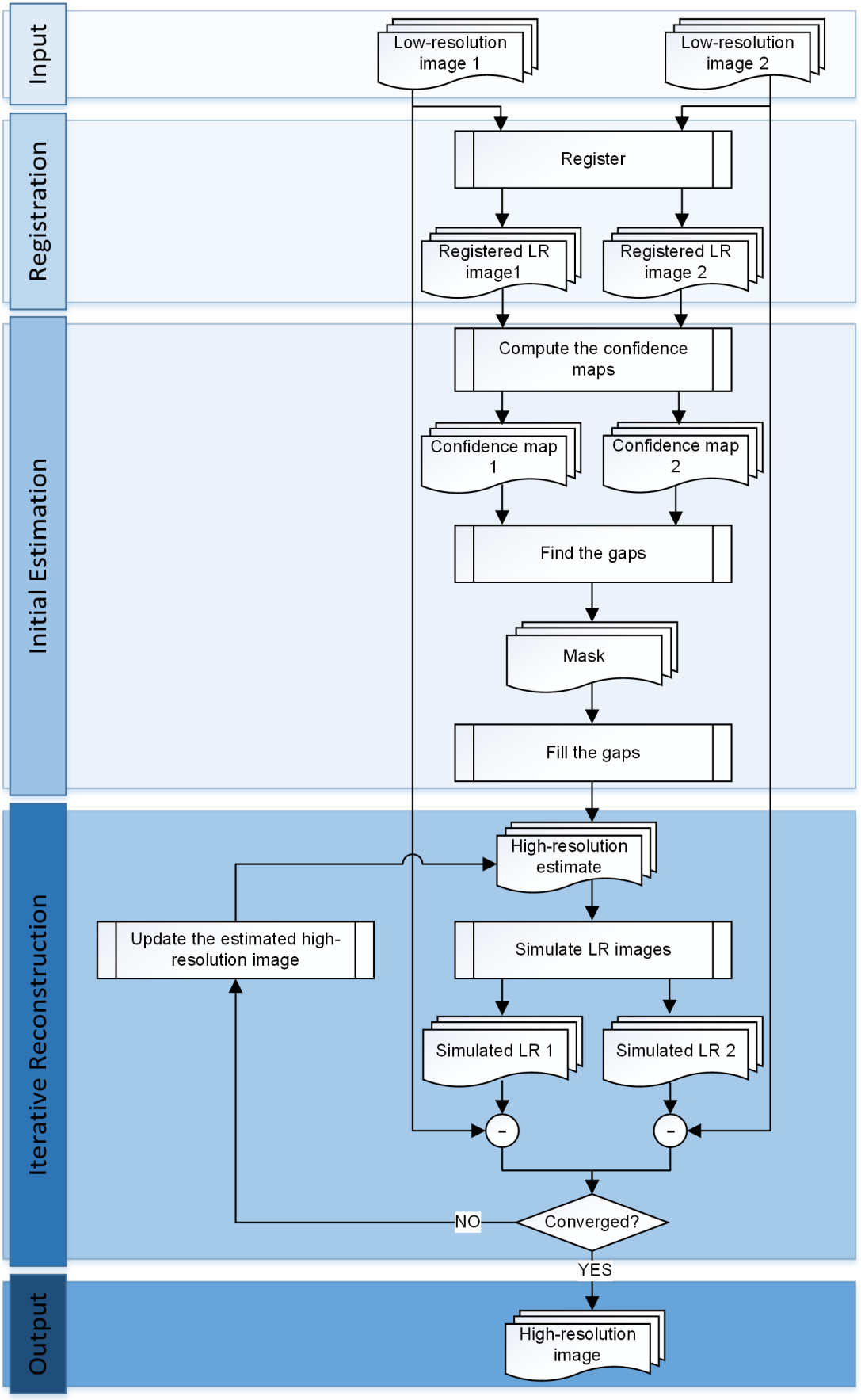
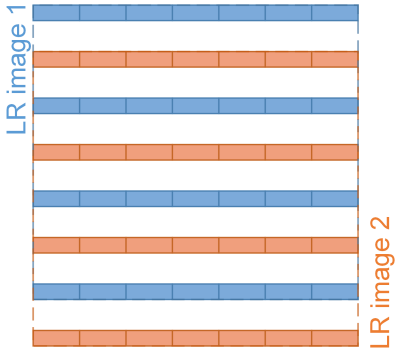


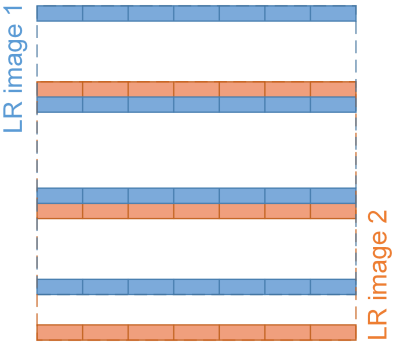
Figure 2.7.: Block diagram of the super-resolution reconstruction algorithm

2.4.1. Acquisition of low-resolution data sets

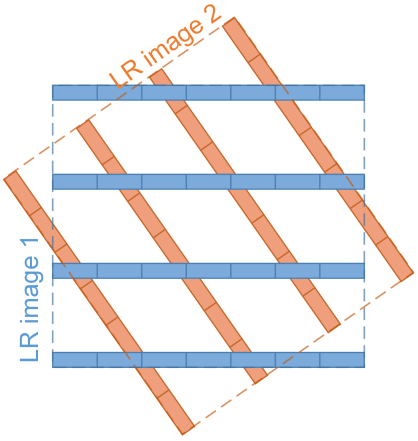
Multi-frame SRR methods assume that the input LR images provide different information about the HR scene. This assumption could be ensured by acquiring LR images with different orientations or offsets, as illustrated in Figure 2.8.



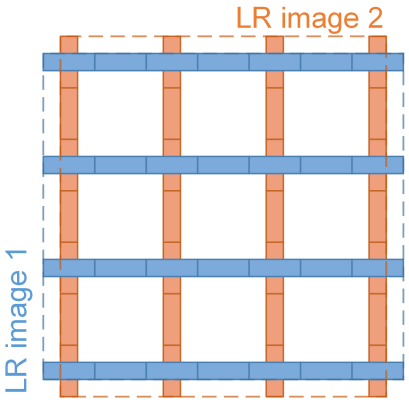
(a) $\alpha_i = 0°$, the LR images have the same longitudinal spacing



(b) $\alpha_i = 0°$, the LR images have different longitudinal spacings



(c) $\alpha_i = 45°$, the LR images have the same longitudinal spacing



(d) $\alpha_i = 90°$, the LR images have the same longitudinal spacing

Figure 2.8.: Possibilities to acquire two LR data sets providing different information about the corresponding HR scene. For simplicity, 2D case is illustrated.

Different offsets. Let α_i be the internal angle between the slice-selection-direction of the first and of the second LR images. If both LR data sets have the same longitudinal direction, i.e. $\alpha_i = 0°$, a translational shift between the first and the second LR images is required to fulfill the mentioned assumption. This shift could be achieved:

- with help of additional hardware to ensure a relative sub-pixel shift, as depicted in Figure 2.8(a)
- or by setting different longitudinal spacings for the first and the second LR images, as illustrated in Figure 2.8(b).

The use of additional hardware is impracticable for clinical application since it imposes additional constrains for the operating room and incurs extra costs. Acquiring LR images with different longitudinal resolution is

not desired since it prevents the use of the highest possible resolution achievable by the scanner (at least for one input data set).

Different Orientations. A better alternative is to increase the angle α_i, i.e. the LR images have different longitudinal directions. The 90° case, shown in Figure 2.8(d), is ideal because it ensures the maximum overlap of the LR images: the HR image can be reconstructed only in this overlapping region $R_\cap$, and therefore it is desirable to maximize this region. By decreasing α_i, $R_\cap$ get smaller till reaching 45° where the overlap is minimal. Further decrease in α_i is associated with an increase in $R_\cap$ which gets again its maximal value for $\alpha_i = 0°$.

For the proposed SRR method it is assumed that the low-resolution images are orthogonal to each other, i.e. $\alpha_i \approx 90°$ (see in Figure 2.9). Although it is possible to reconstruct HR images for any α_i with $90° >> \alpha_i >> 0°$. This case is kept out of focus in this thesis.

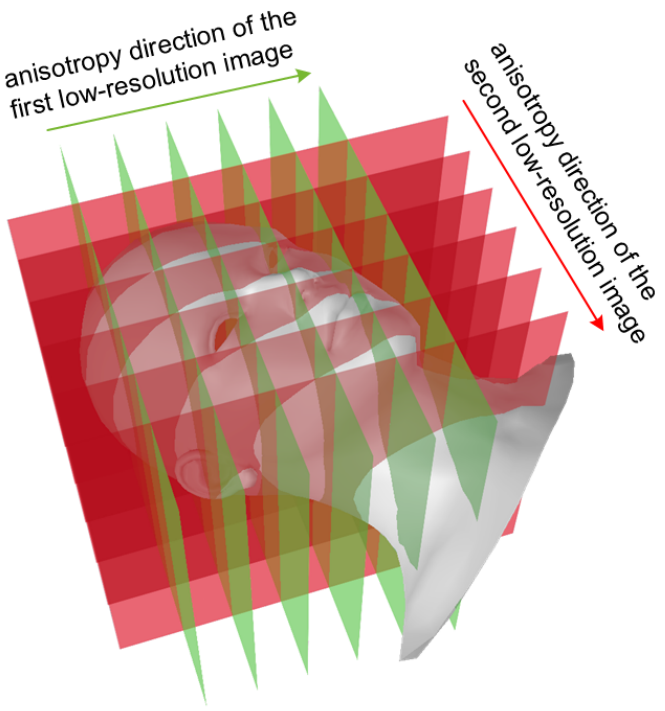


Figure 2.9.: Illustration of the slice-selection direction of the low-resolution images. The acquired CT images have to be more or less orthogonal to each other.

2.4.2. Acquisition model

In order to obtain a meaningful estimate of the HR image, it is critical to accurately model the imaging process. Let $\{Y_n, n = \{1, 2\}\}$ be a set of low-resolution images and X the corresponding unknown high-resolution image. As illustrated in Figure 2.10, the LR data sets can be formulated as a degraded version of the HR image based on the following acquisition model:

$$Y_n = D_n B_n G_n X + v_n \tag{2.19}$$

where G_n is the geometric transformation from the LR frame to the reference HR frame, B_n is the PSF of the acquisition process, D_n is the down-sampling operator, and υ_n is an additive noise. The Equation 2.19 can be simplified by including the geometric transformation, the blur and the down-sampling in a global transformation operator W_n as

$$Y_n = W_n X + \upsilon_n \tag{2.20}$$

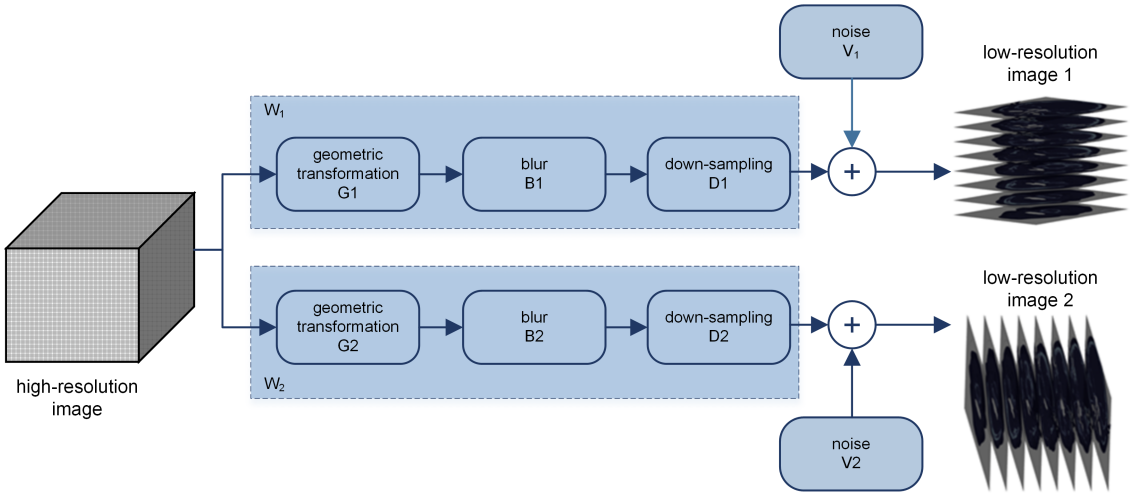


Figure 2.10.: The acquisition model describing the relationship between the unknown high-resolution image and the observed low-resolution images. The observed LR images are regarded as warped, blurred, down-sampled and with noise corrupted versions of the HR image.

2.4.3. Image registration

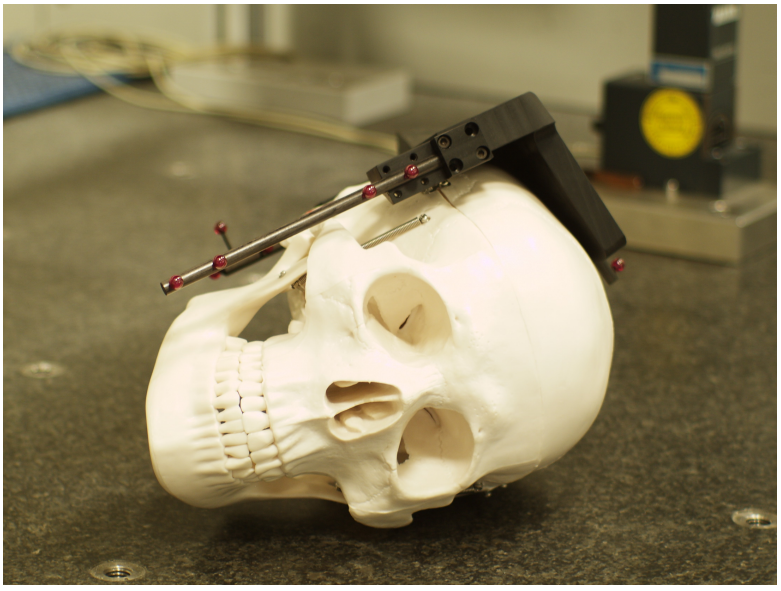


Figure 2.11.: The platform with fiducial markers can be attached to the skull, and used to accurately align the acquired LR CT images to each other.

A key step to apply the SRR is the accurate estimation of the relative position of the LR images with a sub-voxel precision. Achieving a such accuracy is usually not possible, especially in the case of deformable

objects. The region of interest considered in this thesis is the skull. Thus, a rigid registration is sufficient to perform an accurate alignment of the acquired CT images to each other.

The input CT volumes are Hounsfield calibrated, which offers more consistency in the reconstructed values and accordingly lower registration error. In case of isotropic resolution, a simple metric such as the sum of squared differences (SSD) may be sufficient to achieve an accurate registration under the assumption that corresponding points in both images have the same intensity. However, the input LR images have an anisotropic resolution, which may prevent to achieve the desired accuracy since the corresponding voxels have different intensities as they have different anisotropy directions. In order to ensure higher robustness as well as higher accuracy, reference structures (called fiducial markers) are fixed relative to the skull, as depicted in Figure 2.11. The markers remain in the same position relative to the skull during the acquisition of all LR images. Therefore, they can be used to compute the geometrical transformation between the acquired LR volumes.

The LR images are fully automatically registered to each other as following:

1. segment the fiducial markers and localize their center points in both LR CT images, which results in a point set for each LR image
2. determine the point correspondence mapping the elements of the first point set to the elements of the second point set
3. compute the rigid-body transform aligning the LR images to each other

This process assumes that the selected points are invariant relative to the object of interest, i.e. the skull. If the fiducial markers are stuck on the skin, registration error is unavoidable in case of skin motion. In order to avoid such misalignments, the fiducials are fixed to the skull with the help of a ground plate as described in Section 4.1 (see Figure 2.11). Theoretically, 3 markers are sufficient to estimate the (4×4) rigid-body transformation matrix. In this thesis more fiducials, namely 10, are used to define an over determined problem and employ the redundancy to improve the registration accuracy.

The segmentation process of the fiducial markers and the determination of their center points are described in details in Section 4.2.3. The second and third steps of the proposed registration method are solved simultaneously by applying the iterative closest point (ICP) algorithm [BM92]. Thereby the used error metric is the sum of squared distances between the corresponding points. The ICP algorithm aims at determining the rigid-body transformation that minimizes this error.

2.4.4. Confidence-map

It is to be noted that the input data sets (CT images) are not direct measurements of the scanned scene, but rather an indirect reconstruction of it, i.e. based on X-ray projections. Thus, possible degradations are not limited to the acquisition system itself but also influenced by the employed 3D reconstruction algorithm. In order to incorporate knowledge about the several incompleteness sources affecting the data fidelity, confidence maps are introduced. A confidence map describes the fidelity level of each reconstructed CT voxel as a value in the range [0..1], where small values (near 0) refer to high uncertainty, unlike high values (near 1) which mean high fidelity.

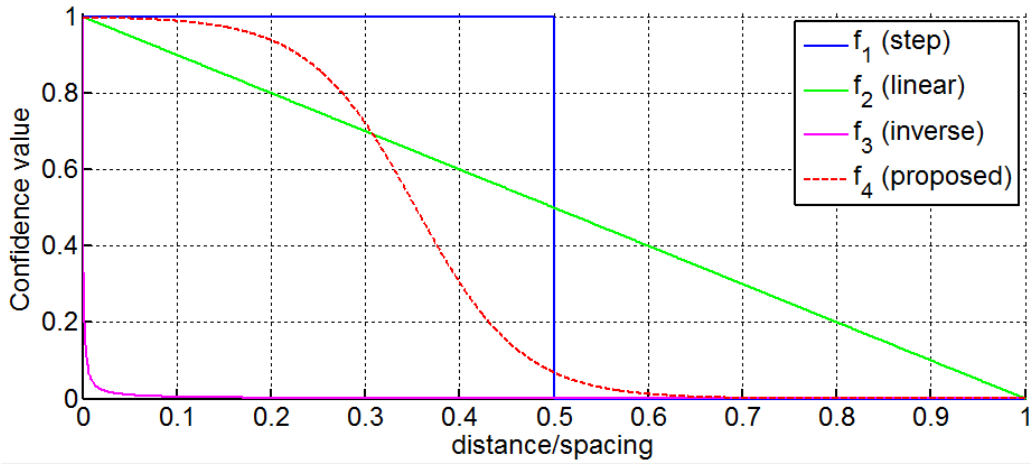


Figure 2.12.: The confidence value as a function of the distance between the aligned low-resolution voxel center and the nearest high-resolution voxel center.

Confidence value as a function of distance. At this level the LR voxels are already aligned to the HR grid. It is assumed that LR voxels in close proximity of a given HR voxel may provide more confident information about this HR voxel. Therefore, a function transforming the distance d_{HL} (between the HR voxel of interest and a given LR voxel) to an appropriate confidence value has to be defined. Accordingly, LR voxels with minimal distances, i.e. $d_{HL} \approx 0$, should have the highest confidence. In contrast, a maximal uncertainty factor should be attributed to distant LR voxels. Several functions can fulfill these requirements, but not all of them are appropriate for the studied SRR problem. In this context, four functions (f_1 to f_4) are investigated: step, linear, Inversely proportional, and Gaussian functions. The corresponding plots are shown in Figure 2.12.

Let s be the HR voxel spacing. The distance weighting functions f_1 to f_4 are defined as following:

- Step function (f_1): only voxels within a radius r_s from the HR voxel of interest, typically $r_s = s/2$, are considered.

$$f_1(d_{HL}) = \begin{cases} 1 & \text{if } d_{HL} \leq r_s, \\ 0 & \text{else.} \end{cases} \tag{2.21}$$

- Linear function (f_2): the uncertainty of a given LR voxel is set proportional to its distance from the HR voxel of interest.

$$f_2(d_{HL}) = \begin{cases} 1 - d_{HL}/s & \text{if } d_{HL} \leq s, \\ 0 & \text{else.} \end{cases} \tag{2.22}$$

- Inversely proportional function (f_3): the confidence value is set Inversely proportional to the distance d_{HL}.

$$f_3(d_{HL}) = \epsilon \cdot \frac{1}{d_{HL} + \epsilon} \tag{2.23}$$

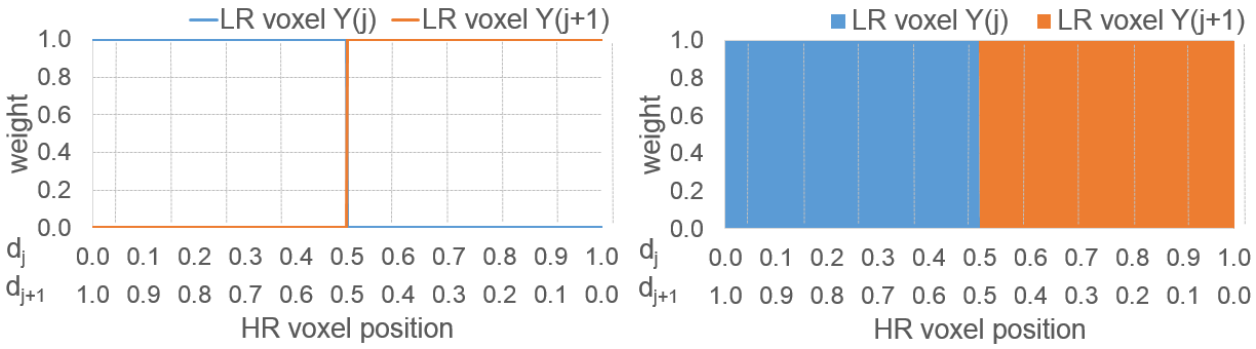
  with ϵ a small value, typically $\epsilon = 10^{-3}$, employed to avoid division by zero.
- Proposed Gaussian function (f_4): this transform follows a Gaussian shape.

$$f_4(d_{HL}) = \frac{1}{1 + exp(2\pi(\frac{2\sqrt{2}d_{HL}}{s} - 1))} \tag{2.24}$$

These functions are studied with respect to two main aspects: the weighting behavior when the LR voxels belong to the same LR image, and when they belong to different data sets. Without loss of generality, the one dimensional case is considered in the following.

Let $Y(j)$ and $Y(j+1)$ be two neighbor points, not necessary belonging to the same LR image, of the HR voxel of interest $X(i)$. The corresponding distances are $d_{(i,j)}$ and $d_{(i,j+1)}$, respectively.

- Step function (f_1): this transformation is very sensitive for minor distance variation if $d_{HL} \approx \frac{s}{2}$. As depicted in Figure 2.13(a), a minor increase in the distance d_{HL} to get larger than $\frac{s}{2}$ results in changing the attributed confidence value from 1 to 0. For instance, if $Y(j)$ and $Y(j+1)$ belong to the same LR image, and the HR voxel $X(i)$ lies in the middle, i.e. $d_{(i,j)} = d_{(i,j+1)} = 0.5$, inconsistency is expected at this position. By increasing d_j, The LR voxel which contributes to $X(i)$ changes discontinuously from $Y(j)$ to $Y(j+1)$, which results in significant distortion and stair-step shaped linear features in the HR image.

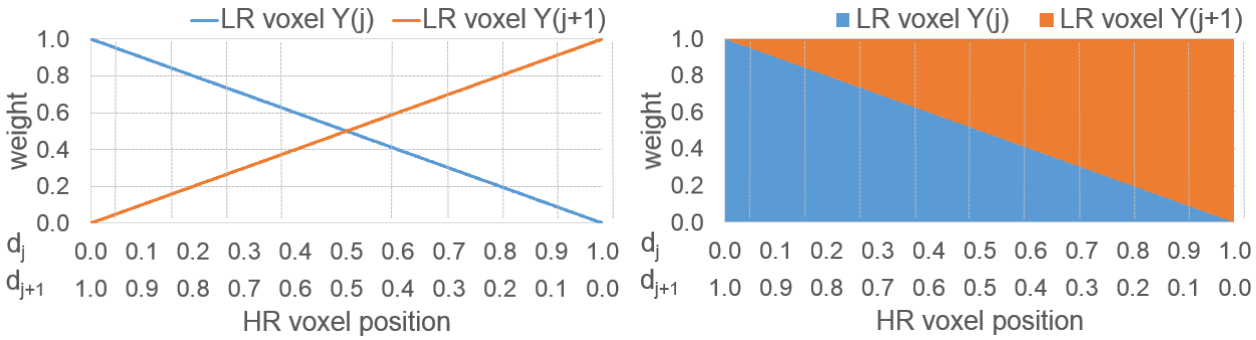


(a) Absolute weights (b) Relative weights

Figure 2.13.: Contribution of LR voxels ($Y(j)$ and $Y(j+1)$) to the HR voxel of interest ($X(i)$) as a step function of the corresponding distances (d_j and d_{j+1}). The distances are normalized relative the HR voxel spacing.

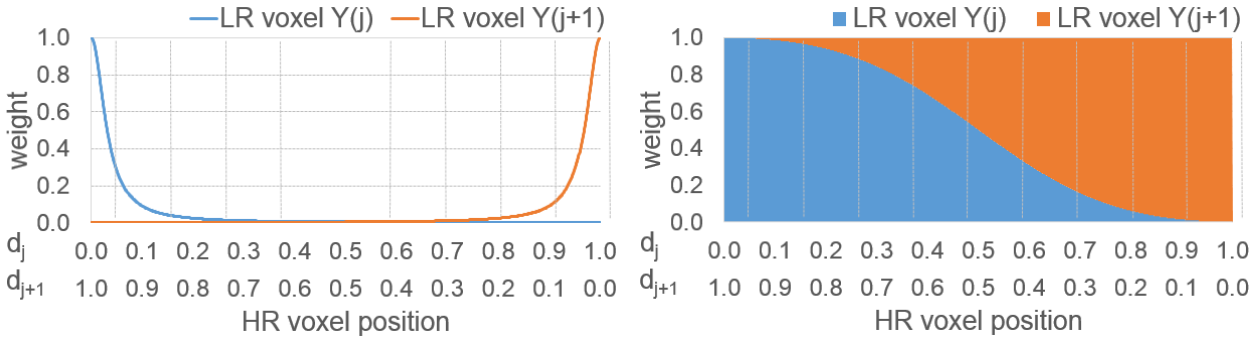
- Linear function (f_2): unlike the step function, no discontinuities are expected in case of linear interpolation. However, the reconstructed image is intended to be smoothed. As shown in Figure 2.14(a), by increasing the distance d_j the contribution of $Y(j)$ decreases. Nonetheless, in the vicinity of the center of $Y(j+1)$, i.e. $d_{j+1} \approx 0$, where $Y(j+1)$ should have exclusive contribution to $X(i)$, $Y(j)$ still have some influence on $X(i)$. For example, if $d_j = 0.1$ and $d_{j+1} = 0.9$, the contributions of $Y(j)$ and $Y(j+1)$ are 0.9 and 0.1, respectively. Ideally, the appropriate contributions should be 1.0 and 0.0, respectively. This over-weighted influence of $Y(j)$ in the proximity of the center of $Y(j+1)$, and vice versa, results in undesired smoothness of the reconstructed HR image.
- Inverse proportional function (f_3): in this case no discontinuities are expected. Moreover, as depicted in Figure 2.15(a), in the proximity of the center of a given LR voxel only the corresponding voxel contributes to the HR voxel of interest. In this transformation the weighting value deceases rapidly with increasing distance, so that the confidence value is maximal only if the LR voxel center is in the vicinity of the HR voxel center, i.e. $d_{HL} \approx 0$. For instance, the confidence decreases to less than 0.5 if the distance $d_{HL}$is larger than 3% of the HR voxel spacing. This characteristic is adverse in case of multiple LR images. For example, let $Y(j)$ and $Y(j+1)$ belong to different LR images and have the distances 0.01 and 0.04, respectively. Despite the minor difference between both distances, i.e. only

3% of the voxels spacing, the confidence attributed to $Y(j)$ is 2.3 times higher compared to $Y(j+1)$. Ideally, $Y(j)$ and $Y(j+1)$ should have similar confidence values.



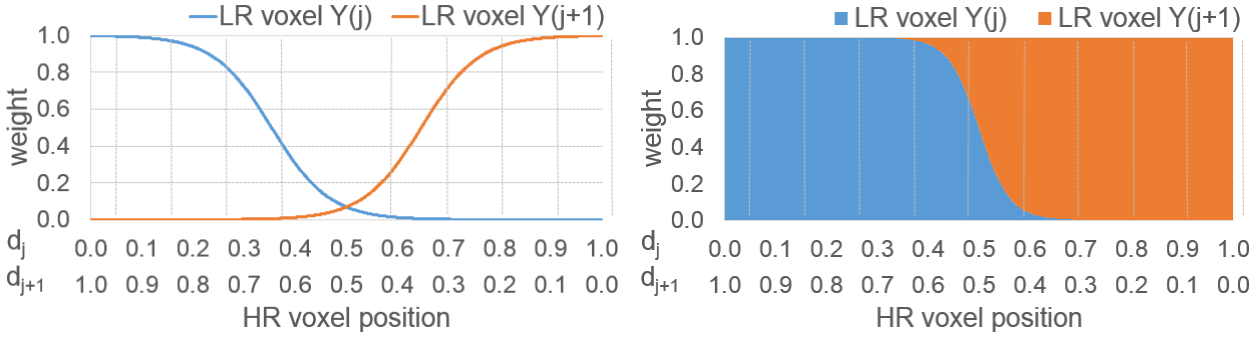
(a) Absolute weights (b) Relative weights

Figure 2.14.: Contribution of LR voxels ($Y(j)$ and $Y(j+1)$) to the HR voxel of interest ($X(i)$) as a linear function of the corresponding distances (d_j and d_{j+1}). The distances are normalized relative the HR voxel spacing.



(a) Absolute weights (b) Relative weights

Figure 2.15.: Contribution of LR voxels ($Y(j)$ and $Y(j+1)$) to the HR voxel of interest ($X(i)$) as a proportional inverse function of the corresponding distances (d_j, and d_{j+1}). The distances are normalized relative the HR voxel spacing.



(a) Absolute weights (b) Relative weights

Figure 2.16.: Contribution of LR voxels ($Y(j)$ and $Y(j+1)$) to the HR voxel of interest ($X(i)$) as Gaussian-shaped function of the corresponding distances (d_j and d_{j+1}). The distances are normalized relative the HR voxel spacing.

- Proposed Gaussian shaped function (f_4): this transformation has no discontinuities, favors LR voxels with small distance d_{HL}, and attributes a weight near to 1 to the LR voxels with small to moderate distances. On the other hand, the weight decreases rapidly in the proximity of the voxels edges, i.e. $d_{HL} \approx 0.5$. For instance, and as depicted in Figure 2.16(b), if the HR voxel lies in the middle of two LR voxels belonging the the same LR image, these later have the same weight, i.e. 0.5.

Due to the adequate behavior of the proposed distance-weighting function ($f4$), this transformation is employed to compute the confidence term $\mathcal{C}_D$.

Let $d(i, j, n)$ be the distance between the center of the nearest HR grid cell $X(i)$ at position i and the center of a given LR voxel $Y_n(j)$ of the n-th LR image at the position j. The partial confidence term $\mathcal{C}_D(i, j, n)$ penalizes the voxels with large distance from the HR voxel of interest by decreasing their fidelity level following this equation

$$\mathcal{C}_D(i, j, n) = \frac{1}{1 + exp(2\pi(\frac{2\sqrt{2}d(i,j,n)}{s_{in}} - 1))} \tag{2.25}$$

where s_{in} is the in-plane spacing.

Confidence value as a function of noise. CT scans are typically affected by Poisson noise. By increasing the noise level, low contrast objects, such as soft tissue, may be obscured unlike high contrast objects which remain visible. Since the input data are not direct measurements of the 3D scene, the noise present in the 3D reconstruction depends on several factors: the X-ray tube voltage, the radiation intensity, the section width, the z-filtering and the reconstruction process [TA98].

In case of Filtered Back-Projection (FBP) algorithm, which is the standard reconstruction method on most of scanners, the standard deviation in Hounsfield units (HU) due to Poisson noise is proportional to $\sqrt{slicethikness \times mAs}$ [Gol07]. This means that the Poisson noise can be decreased by increasing the mAs or the slice thickness. Increasing the mAs would increase the patient dose, thus a tradeoff between noise, resolution, and radiation exposure should be found. A drawback of the filtered back-projection approach is that the projection data are assumed to be accurate, and therefore ignores that low photons counts may result in large Poisson error. In contrast, iterative reconstruction approaches take account of this aspect by assuming that smoother images are more probable. Theses methods can substantially reduce noise and improve image quality even at lower radiation dose. Thus, the relationship between noise and dose is different compared to FBP.

Noise variation of the input LR images can be evaluated using head and body sized water phantoms. In ideal case, two voxels of identical material would result in identical intensities. This property serves to estimate the noise level by computing the standard deviation of a homogeneous area in the CT reconstructions. For illustration purposes, the variation of noise level by changing the scan parameters is summarized in Table 2.2. Thereby relative noise values are computed in order to normalize the noise values for different phantoms and scans parameters.

The partial confidence term $\mathcal{C}_\upsilon$ computed based on the noise level is defined as

$$\mathcal{C}_\upsilon(Y_n) = 1 - \alpha_\upsilon \left(1 - \frac{\upsilon_n}{\max_m(\upsilon_m)}\right) \tag{2.26}$$

where υ_n is the noise level of the n-th LR image Y_n normalized by dividing it through the maximum noise level over all LR images ($\max_m(\upsilon_m)$). The factor α_υ ($\in [0, 1]$) weights the contribution of the relative noise on the final confidence value $\mathcal{C}$. If the input LR images are acquired with the same scanner and parameters, this confidence term $\mathcal{C}_\upsilon$ has no influence on $\mathcal{C}$.

	Parameter Setting	Relative noise
Tube voltage (kV) (Head)	80	1.89
	100	1.31
	120	**1.00**
Tube voltage (kV) (Body)	80	2.48
	100	1.40
	120	**1.00**
	140	0.79
Tube current (mA)	33	2.65
	80	1.88
	160	1.37
	320	**1.00**
	400	0.91
Scan time (s)	0.5	1.33
	1.0	**1.00**
Slice thickness (mm), Collimation (# $\times$ mm)	5 (1 $\times$ 5)	1.03
	10 (1 $\times$ 10)	0.75
	3 (20 $\times$ 0.6)	1.22
	6 (20 $\times$ 0.6)	0.88
	12 (20 $\times$ 0.6)	0.64
	2.4 (24 $\times$ 1.2)	1.37
	4.8 (24 $\times$ 1.2)	**1.00**
	7.2 (24 $\times$ 1.2)	0.83
	9.6 (24 $\times$ 1.2)	0.73
	14.4 (24 $\times$ 1.2)	0.61

Table 2.2.: Image noise as function of scan parameters [NDMS05]. The standard settings to which the other values are normalized are set in bold.

Confidence value. The confidence value is computed for each voxels as the multiplication of the partial confidence terms:

$$\mathcal{C}(i, j, n) = \mathcal{C}_D(i, j, n) \cdot \mathcal{C}_\upsilon(Y_n) \tag{2.27}$$

2.4.5. Initial super-resolution estimate

Fusing reference voxels

The initialization of the SR volume does not influence the stability, but might influence the convergence speed and the solution to which the method converge. In general, the better the initial estimate matches the final HR images, the faster the process converges towards a stable solution. Usually, the average image of the two LR data sets is chosen as a reasonable initial guess. This initialization is not recommended in case of highly anisotropic voxel since it would result in high smoothness degree. This thesis proposes to

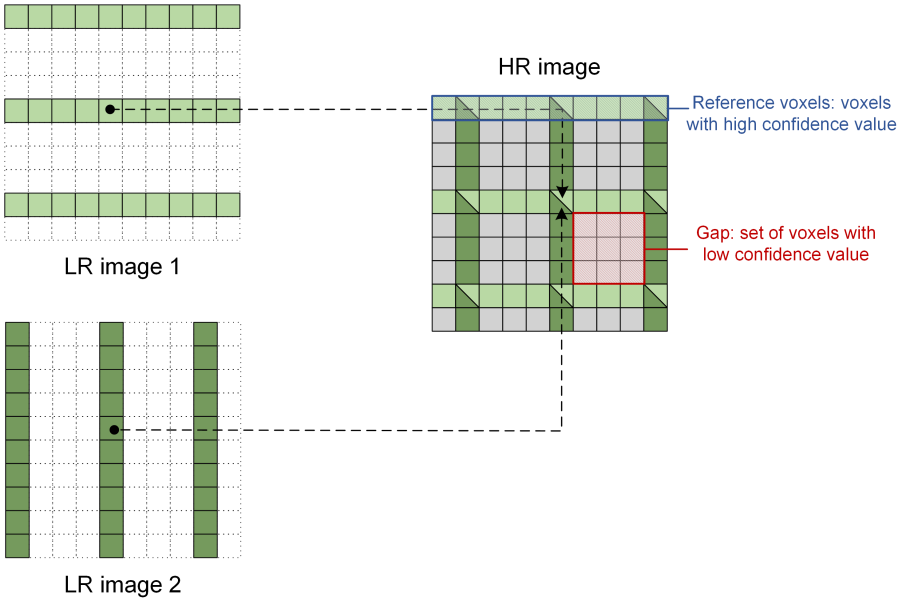


Figure 2.17.: Illustration showing the fusing of the low-resolution images in a high-resolution grid, and the resulting regular pattern of gaps.

fuse the input data sets based on the confidence values. For instance, if a given input LR voxels has a high confidence value (≈ 1) and the voxel from the second LR image corresponding to the same position has lower confidence value (like 0.2) then the corresponding HR voxel value to be recovered should be mainly influenced by the first voxel since it is more reliable to offer more accurate information about the HR voxel of interest. As illustrated in Figure 2.17, each voxel $X^0(i)$ of the initial SR estimate X^0 at the position i is computed as the weighted sum of all LR voxels $Y_n(j)$ multiplied by the corresponding confidence values $C(i,j,n)$:

$$X^0(i) = \frac{1}{\sum_n \sum_j C(i,j,n)} \sum_n \sum_j C(i,j,n) \cdot Y_n(j) \tag{2.28}$$

Recovering missing voxels

In case of highly anisotropic data sets, fusing the LR images based on the Equation 2.28 results in several gaps at positions where no confident input data is present. Given that the acquired LR images are orthogonal as proposed in Section 2.4.1, the gaps should occur with a regular pattern, in ideal case as depicted in Figure 2.18. These gaps can be easily detected by extracting voxels which have low confidence values.

Filling the gaps by employing a simple interpolation, such as linear interpolation, results typically in smooth reconstruction. In order to fill the detected gaps in a “meaningful” way while keeping sharp details, an inpainting approach should be applied. Most of inpainting algorithms are classified into two groups: geometry- and texture-oriented approaches. Geometry-oriented techniques try to continue the geometric structures present in the image, such as level lines and edges [YTZ*09]. The texture-oriented methods, known also as exemplar-based techniques, estimate missing values by copying learned patches from sample images [AFCS11]. For further reading, the reader is referred to the publication of Guillemot et al. [GL14].

Due to the nature of the investigated CT images, geometry-oriented methods are more appropriate: the organs contours are typically sharp and change continuously between adjacent image slices. In addition, such methods do not require training data and have relatively low computational cost.

In this thesis, a 2D direction-oriented interpolation inpainting method [YTZ*09] is extended to fill 3D gaps. The fundamental idea of this technique is based on the anisotropic filtering theory [Wei97, AFP91]: the direction of geometric structures, like edges, is detected (or approximated) and then a directional interpolation is performed along the estimated edges. This idea can be extended to 3D by performing the interpolation also along surfaces.

An alternative is to apply the recovering (interpolation) slice-wise. As illustrated in Figure 2.18, the combined 3D image consists of a stack of slices with isotropic pixels. By defining the slice-selection-direction as the direction orthogonal to the out-of-plane orientations of the input LR images, the gaps occur with a regular patter along the said direction. Filling the gaps in all of these slices results in recovering all missing voxels of the isotropic HR image. Following this scheme, the 3D inpainting can be reduced to 2D inpainting.

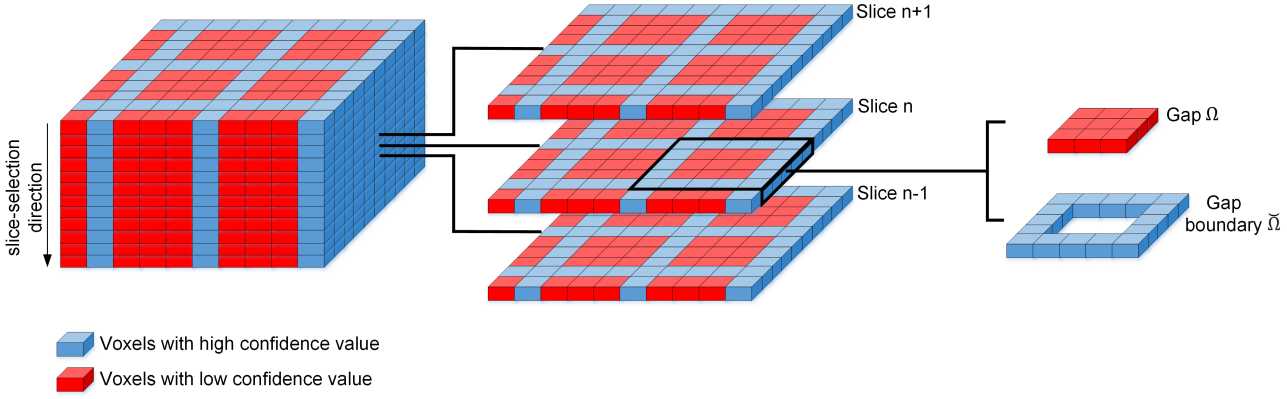


Figure 2.18.: Illustration showing the regular pattern of gaps reaming after fusing two orthogonal LR images.

The scanned objects (e.g. head) have definite geometric structures in the space, such as in the case of edges or surfaces. The structure continuity should be preserved while performing the inpainting step e.g. by interpolating along the edges/surfaces, not across them. A key processing step is therefore the accurate estimation of the main interpolation direction. In this context a robust solution is proposed to determine the main (prominent) variation direction at missing voxels. Accordingly, we introduce a space variant interpolation algorithm which steers with an elliptic variation of a Gaussian distribution as a filtering kernel to a specific interpolation direction.

Given a gap region Ω, the corresponding gap boundary $\breve{\Omega}$ is defined as the set of reference voxels (voxels with high confidence) surrounding the said gap, as illustrated in Figure 2.18. The Boundary voxels $\breve{\Omega}$ are used as sampling points to interpolate the free space in between (slice-wise), i.e in 2D space. In order to determine the appropriate interpolation direction for each gap, we proceed as follows:

1. calculate a global (representative) 3D gradient vector for the gap Ω based on the 3D gradient vectors along the gap boundary $\breve{\Omega}$. For each boundary voxel the gradient vector is computed by applying a 3D-Sobel operator to achieve a better structure continuity in 3D space. Thus, also the neighbor slices of the current slice, i.e. the previous and the next slices as illustrated in Figure 2.18, are considered.

2. compute the covariance matrix $\Sigma_{\breve{\Omega}}$ of the gradient vectors calculated in the first step, more precisely, from the projection of the said 3D gradient vectors on the interpolation plane, since the interpolation is performed there. The eigenvalues of $\Sigma_{\breve{\Omega}}$ represent an approximation of the gradient vector distribution, whose major axis defines the main variation direction.

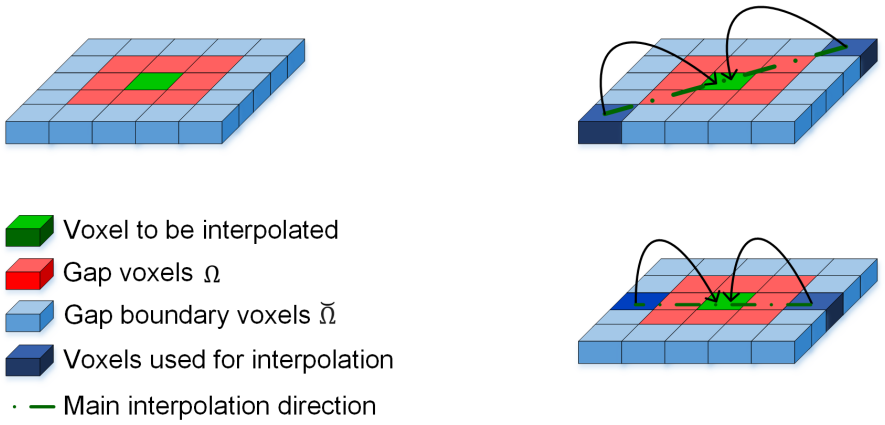


Figure 2.19.: Illustration showing the boundary voxels employed to recover the value of the voxel to be interpolated. The selection of the boundary voxels with main contribution to the interpolated voxel depends on the estimated interpolation direction.

It is to be noted that not all gap regions have a prominent variation direction, such as the case in homogeneous regions. At this level it is to be distinguished between two main cases:

- the gap contains one and only one (unique) prominent variation direction, e.g. in case of a line-shaped structure. In such a case only the voxels which lie on this structure should be employed to interpolate the voxels in between (which lie also on the said structure). Therefore, mainly the voxels which are on the line orthogonal to the dominant variation direction direction and crossing the voxel to be interpolated, should be considered in the interpolation (see Figure 2.19).
- no unique prominent variation direction is detected. In such a case all current boundary voxels should contribute to the interpolation based on their distances from the voxel to be interpolated.

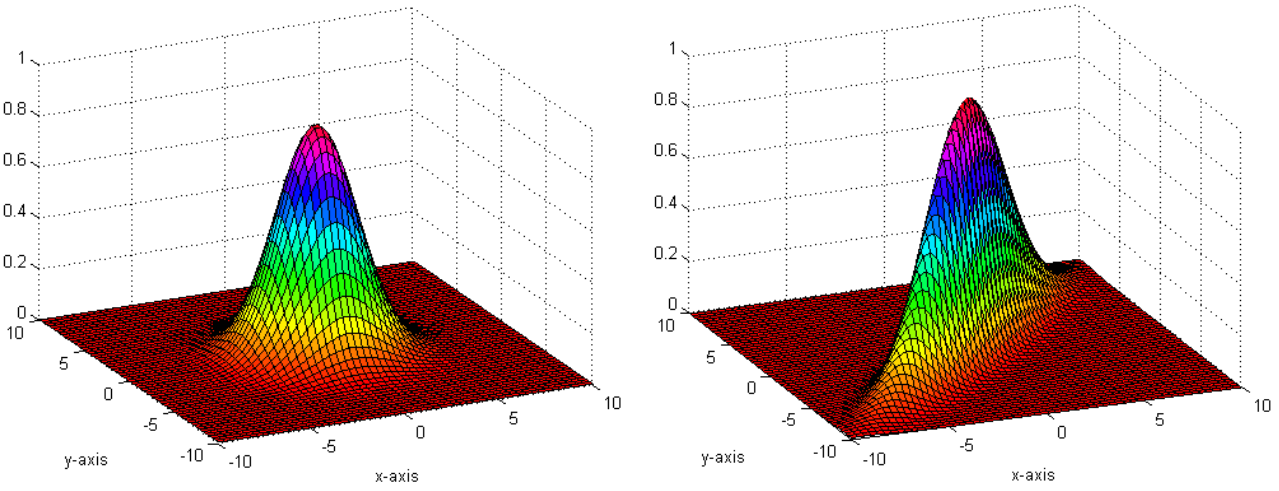
In order to distinguish between these two cases, a metric has to be defined. For this purpose a stretching factor σ ($\sigma \geq 1$) is introduced. It is defined as the ratio of the largest eigenvalue λ_{max} of the covariance matrix $\Sigma_{\breve{\Omega}}$ to the smallest eigenvalue λ_{min}, i.e $\sigma = \lambda_{max}/\lambda_{min}$. The factor σ would tend to 1 in case of homogeneous regions and would increase ($\gg 1$) if a prominent variation direction is detected. For instance, in case of line-like structures, one eigenvalue is much larger than the other eigenvalues. Accordingly, the behavior of the interpolation kernel is defined as follows:

- if $\sigma \approx 1$, the kernel results in a structure-independent isotropic weighting as illustrated in Figure 2.20(a)
- if $\sigma \gg 1$, the kernel follows a discus shaped propagation as depicted in Figure 2.20(b)

An appropriate interpolation kernel $\kappa(i,j)$ is proposed to fulfill this behavior. It consists of three terms: distance d, structure s, and penalty terms p (see below). The kernel $\kappa(i,j)$ creates weights for the boundary voxels based on this equation

$$X^1(i) = \frac{\sum_{j \in \breve{\Omega}} X^0(j) \overbrace{d(i,j)\, s(i,j)\, p(j)}^{\text{interpolation kernel } \kappa(i,j)}}{\sum_{j \in \breve{\Omega}} \kappa(i,j)} \tag{2.29}$$

where $X^1(i)$ is the intensity of the estimated SR voxel at the interpolation position i and $X^0(j)$ is the intensity of the boundary voxel at position j.



(a) Isotropic Gaussian kernel ($\sigma = 1, \mu = 2$)

(b) Directional anisotropic Gaussian kernel ($\sigma = 5, \mu = 2$)

Figure 2.20.: Illustration of isotropic and anisotropic Gaussian kernels.

Distance weighting. It is a non-directional weighting function declining over distance $d(i, j)$ which is the Euclidean distance between the voxels at positions i and j. This weighting term favors boundary voxels which are closer to the interpolated voxel and penalizes distant voxels:

$$d(i, j) = e^{-\frac{d^2(i,j)}{2\mu^2}} \tag{2.30}$$

where the factor μ determines the smoothness of the Gaussian kernel.

Structure function. It aims at performing the interpolation along the main structure direction, i.e. orthogonal to the global gradient vector $\vec{g}_j$ of the actual gap. Let $\vec{d}_{ij}$ be the vector between the current boundary voxel $X(i)$ and the interpolated voxel $X(j)$, as illustrated in Figure 2.21(a). If $\vec{d}_{ij}$ is orthogonal to $\vec{g}_j$ (Fig. 2.21(b)), the corresponding voxel $X(i)$ should contribute to $X(j)$. In contrast, if $\vec{d}_{ij}$ and $\vec{g}_j$ are collinear (Fig. 2.21(c)), no contribution should be reported. For this purpose the internal angle θ between $\vec{g}_j$ and $\vec{d}_{ij}$ is employed: the contribution is set proportional to $\sin(\theta)$.

On the other hand, a small gap region does not necessarily contain a unique sharp structure ($\sigma \approx 1$). For example in case of homogeneous regions any interpolation direction can be chosen, since all boundary voxels have more or less the same value. However, it is more appropriate to not favor a specific interpolation direction and consider all boundary voxels based on their distance from the voxels to be recovered.

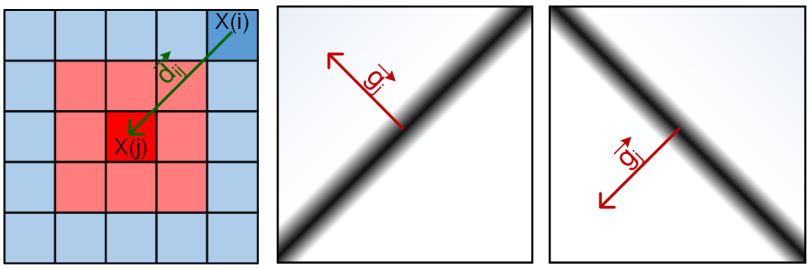


Figure 2.21.: Illustration of the vector $\vec{d}_{ij}$ between the current boundary voxel $X(i)$ and the interpolated voxel $X(j)$ (a). Examples in which $\vec{d}_{ij}$ is orthogonal or parallel to the main variation direction are sown in (b) and (c), respectively. In the case shown in (b) the contribution of $X(i)$ to $X(j)$ is maximal, unlike the case shown in (c).

The stretching factor σ ($\sigma \geq 1$) is employed to determine the spatial expansion of the interpolation kernel. In case of homogeneous regions, ($\sigma \approx 1$) $s(i,j)$ tends to 1, so that a structure-independent weighting is applied. If a prominent variation direction is detected ($\sigma \gg 1$), the interpolation kernel is compressed along this direction by $\frac{1}{\sigma}$ and stretched along the orthogonal direction by factor σ. Now based on the inner angle θ between $\vec{d}_{ij}$ and $\vec{g}_j$, the contribution of the voxel $X(i)$ is defined as follows

$$s(i,j) = s(\theta) = \sqrt{\frac{1}{\sigma^2 \cos^2(\theta) + \frac{1}{\sigma^2}\sin^2(\theta)}} \tag{2.31}$$

The structure term $s(i,j)$ is maximal for $\theta = 90^\circ$ ($\sin(\theta) = 1$), and minimal for $\theta = 0^\circ$ ($\sin(\theta) = 0$). For instance, if $\sigma = 5$, the contribution is equal to 5 for $\theta = 90^\circ$ and equal to 0.2 for $\theta = 0^\circ$. If $\sigma = 1$, the contribution is always the same (=1), independently of θ.

Penalty function. The gradient vectors of the actual boundary voxels ($\subset \breve{\Omega}$) are ideally intended to have similar directions. Therefore, a penalty function $p(j)$ is introduced to measure the gradient quality of each boundary voxel and penalize misleading ones. The similarity between the gradient vector $\vec{g}_j$ at position j and the global gradient vector $\vec{g}_{global}$ of $\breve{\Omega}$ is defined as

$$p(j) = \frac{|\vec{g}_{global} \cdot \vec{g}_j|}{\|\vec{g}_{global}\| \|\vec{g}_j\|} \tag{2.32}$$

If $\vec{g}_j$ and $\vec{g}_{global}$ are collinear, p(j) is set to one. Otherwise $p(j)$ tends to zero if both vectors get orthogonal to each other.

2.4.6. Iterative reconstruction

After performing an initial estimation for the SR volume, a set of LR images $\widehat{Y}_n$ is simulated based of the acquisition model described by Equation 2.19 (see Section 2.4.2). If the estimate SR image were the correct HR volume, then the simulated LR images have to be identical to the acquired LR images (under the assumption of accurate imaging model). The initial estimate is improved by back-projecting the

difference between the simulated and observed LR images. The iterative reconstruction is terminated if the said difference is larger than in the previous iteration, or if the difference is less than a threshold (10^{-3}).

Mathematically, the proposed iterative back-projection approach is applied to minimize the following error function:

$$\epsilon^{(k)} = \sqrt{\sum_n \sum_j \mathcal{C}(i,j,n)(Y_n(j) - \widehat{Y}_n^{(k)}(j))^2} \tag{2.33}$$

where k is the iteration number. Unlike classical approaches, the confidence term is introduced in the error function in order to favor against noise and missing data. An update of the reconstructed volume is then performed based on the following equation

$$\widehat{X}^{(k+1)}(x) = \widehat{X}^{(k)}(i) + \lambda \sum_{j \in \cup_n \Delta_{n,i}} \mathcal{C}(i,j,n)\Big(Y_n(j) - \widehat{Y}_n^{(k)}(j)\Big) \times h_{ij}^{BP} \tag{2.34}$$

where λ is the step length, h^{BP} is the back-projection kernel and $\Delta_{n,i}$ is the set of the LR voxels (of the n-th LR data set) influenced by the the HR voxel at position i. The BP kernel h_{ij}^{BP} weights the contribution of the LR voxels to update the correspondent HR voxels. The term $\mathcal{C}(i,j,n)$ serves to incorporate precomputed confidence values. A total variation regularization is applied in gaps regions [CS05].

2.5. Evaluation and results

In this section the performance of the proposed reconstruction algorithm is compared to other super-resolution reconstruction techniques. Experiments are conducted on simulations and real patient data. The reconstruction of high-resolution images were performed using the following methods:

- B-Spline interpolation (B-Spline)
- Iterative back-projection SRR (IBP)
- Robust M-estimation SRR as proposed in [RKS*10b] [GEW10] and integrated in BTK-toolkit [ROP*13] (BTK)
- The proposed confidence-map based SRR (cMap)

Error measure. For quantitative analysis, comparison is performed in term of peak signal to noise ratio (PSNR):

$$PSNR = 20 log_{10} \cdot \frac{MAX}{\sqrt{MSE}} \tag{2.35}$$

where MAX is the maximum possible voxel value of the HR image, and MSE is the mean square error. The error is defined as the difference in intensities between the reconstructed image and the corresponding reference volume. A higher PSNR indicates higher reconstruction quality.

Phantoms. In order to study the influence of SRR on the segmentation accuracy, these calibrated phantoms have been scanned:

- step-cylinder phantom made of bone equivalent material
- double-ball-bar phantom made of synthetic rubin and carbon fiber

Material	bone equivalent
No. steps	5
Height [*mm*]	80
Diameter [*mm*]	20 / 40 / 60 / 80 / 100

Step-cylinder phantom

Material ball	synthetic rubin
Material bar	carbon fibre
Length [*mm*]	30 / 90
Diameter [*mm*]	5 / 10

Double-ball-bar phantom

Figure 2.24.: Phantoms used to evaluate the reconstruction quality.

Figure 2.24 shows and summarizes the specification of the used phantoms. The step-cylinder phantom has been made from a bone equivalent material and sized to respect typical radiation conditions (attenuation, length) in the cranial area [Seb14].

Since the proposed SRR method was motivated by he lateral-skull surgery, a cadaver skull was scanned with an industrial CBCT scanner in order to study the reconstruction quality in the temporal bone region (see Figure 2.25).

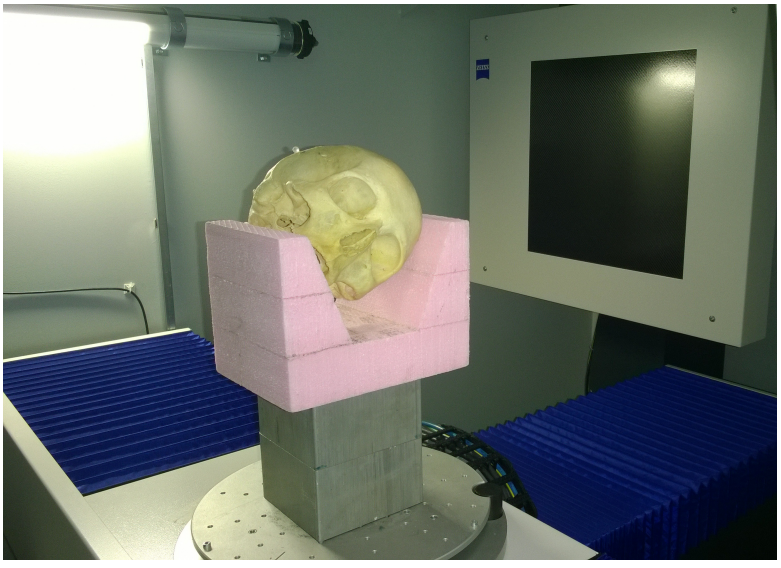


Figure 2.25.: A cadaver skull scanned with an industrial CBCT scanner.

Data acquisition. 3D images were acquired for all three phantoms with an industrial CBCT scanner in high resolution mode. The voxel dimensions are $0.253 \times 0.253 \times 0.253$ mm^3 . For each CT volume, 1200 projections over 360° were acquired. The high X-ray dose (200 kV, 0.3 mA) was employed to achieve

better image quality. The acquired 3D images serve as "ground truth", since industrial CT scanners have lower uncertainty in comparison to current medical CT systems as shown in [RSM12]. Further evaluations were performed based on CT images acquired with a medical spiral CT scanner (SOMATOM Sensation, Siemens, Germany).

2.5.1. Registration accuracy

The reconstruction quality is sensitive to excessive registration error. It was assumed that the registration process can achieve sub-voxel accuracy. The fulfillment of this assumption was quantitatively evaluated by calculating the target registration error (TRE) which is the Euclidean distance between the position of a target mapped by the estimated transformation and the reference transformation. Since the "ground-truth" transformation is unknown, LR images were simulated from HR volumes and then initial transformations were generated by perturbing the reference transformation (employed to simulate the LR images) by adding randomly generated rotations and translations. Twelve HR volumes were used to generate the LR images. For each HR volume, 100 LR data sets were created with different randomly generated initial deviations.

Two registration approaches were compared in term of registration accuracy, namely the proposed fiducial-based registration and an intensity-based registration (combined with mutual information metric (MI) and Powell's optimizer). As initialization two cases are distinguished: initial small deviation (TRE $\leq$ 5mm) and large initial deviation (TRE $>$ 5mm). The results are summarized in Tables 2.3.

	Intensity-based registration		Proposed fiducial-based registration	
initial error [mm]	$\leq$5	$>$5	$\leq$5	$>$5
mean	0.38	3.38	0.13	0.14
std	0.16	1.21	0.07	0.07
max	0.83	impossible	0.25	0.26

Table 2.3.: Target registration error after applying an intensity-based registration and the proposed fiducial-based registration. The TRE is scaled relative to the high-resolution voxel spacing.

In case of small initial deviations ($\leq$5mm) the intensity-based registration was able to achieve sub-voxel accuracy. However, this approach failed if the initial deviation is too large. Thereby, the registration is considered as unsuccessful if the TRE exceeds the specified sub-voxel accuracy. Intensity-based registration methods are susceptible to local optimum in the cost function. Thus, initial transformation has to be close to the correct transformation for a reliable convergence. In case of large initial deviation ($>$5mm) the intensity-based method may stuck in local optimum, which results in large TRE.

A better behavior was achieved with the proposed fiducial-based registration. It was able to converge award the right optimum for small as well as for large initial deviations. Thereby the TRE was three times smaller compared to the intensity-based registration. It is well known that in case of isotropic resolution, intensity-based registration is expected to be more accurate than feature-based approaches [PDBF12]. This expectation is not satisfied in case of high anisotropy due to discrepancies in image intensities because the LR images have different anisotropy directions. Since this differences in intensities are more or less random, the MI metric (as well as all other common metrics) can not account for these discrepancies, which results in relatively high registration error. In contrast, the anisotropy aspect has a limited influence

on the fiducial-based registration if the fiducial markers expand over several slices. Since their ideal spherical shape is known, it is relatively easy to accurately estimate their center points. Consequently, higher registration accuracy could be achieved.

Another advantage of the fiducial-based registration compared to the intensity-based approach is the short processing time. It requires less than 0.19 seconds, so that more than 50 times faster than the intensity-based approach which takes ca. 9.7 seconds (on GPU).

2.5.2. Reconstruction quality

2.5.2.1. Simulated data

Twelve data sets of four cadaver skulls were acquired with an industrial CBCT scanner in high resolution mode. Each skull was scanned three times with different orientations. The acquired images serve as ground-truth for the simulated LR data sets. Low-resolution volumes were simulated by down-sampling the acquired 3D images along two orthogonal directions and then adding noise to them.

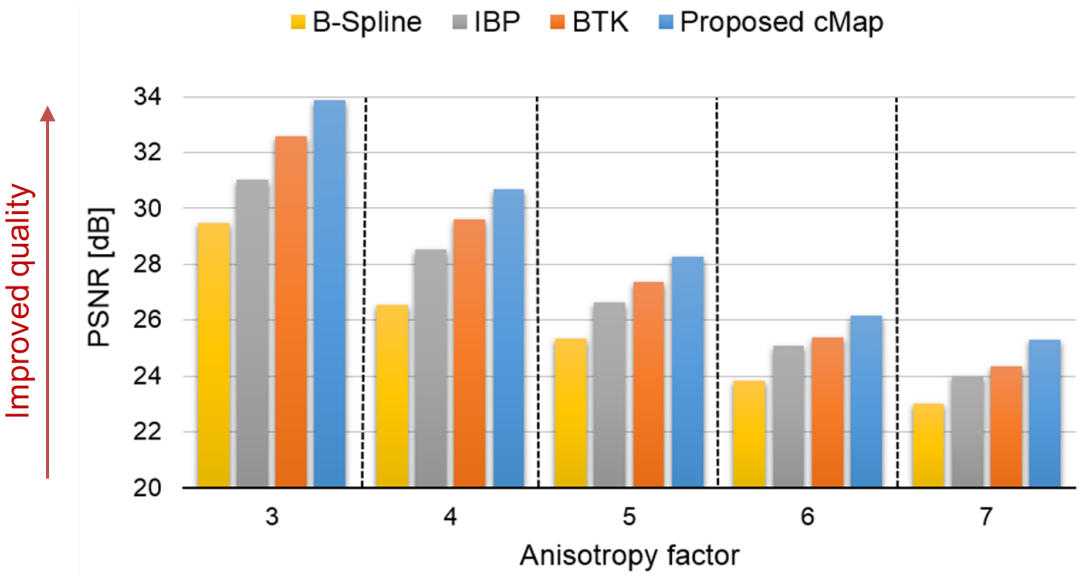


Figure 2.26.: PSNR of high-resolution images reconstructed with different SRR approaches. Several anisotropy factors (3 to 7) are evaluated.

Robustness against high anisotropy. 3D images acquired with spiral CT have a typical slice thickness which is approximately 3 to 7 times bigger than the in-plane spacing. In order to evaluate the ability to reconstruct accurate HR volumes for different anisotropy factors, several SRR methods were applied to LR data sets simulated with different slice thicknesses. As depicted in Figure 2.28, it can be seen that by increasing the anisotropy factor the quality of reconstructed image became lower and fine structures get lost. This observation correlates with the corresponding quantitative results shown in Figure 2.26.

The peak signal to noise ratio (PSNR) was computed over all data sets for the different investigated SRR methods. The corresponding results, reported in Figure 2.27, indicate that the proposed cMap approach yields significantly lower PSNR compared to the other methods.

A challenging case is the reconstruction of HR volume from LR images with large slice thickness. Even in case of high anisotropy factor (=7), the quality of the reconstructed images with the proposed confidence-map based SRR remains acceptable. For qualitatively good results, we suggest to keep the anisotropy factor $\leq$ 5. This suggestion is supported by the quantitative evaluation of the segmentation accuracy presented in Section 2.5.2.1. For illustration purposes, reconstruction results in case of anisotropy factor = 5 are shown in Figure 2.29. The ground truth image (acquired with an industrial CT scanner) and a corresponding LR image are shown in the left column. The results obtained by the different reconstruction algorithms are shown in the second and third columns. It can be seen that the proposed cMap-based SRR can recover sharper edges and more details compared to the other approaches.

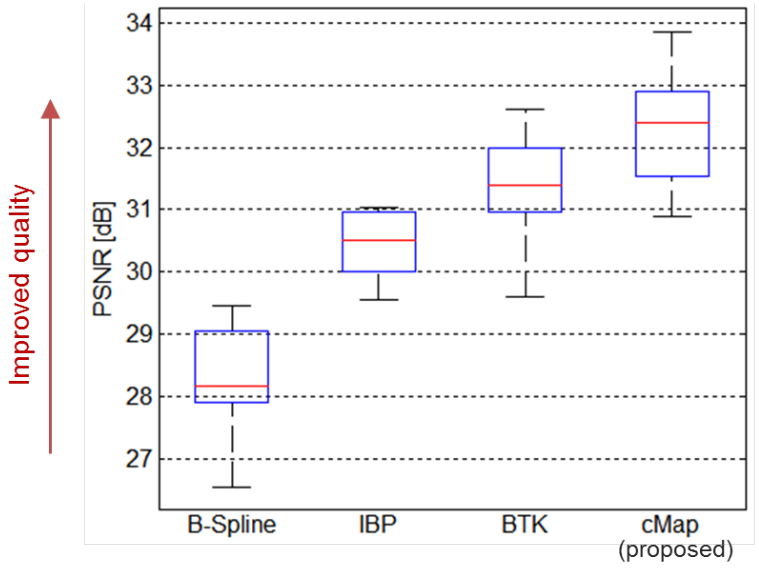


Figure 2.27.: PSNR of high-resolution images reconstructed with different SRR approaches. The anisotropy factor is equal to 3.

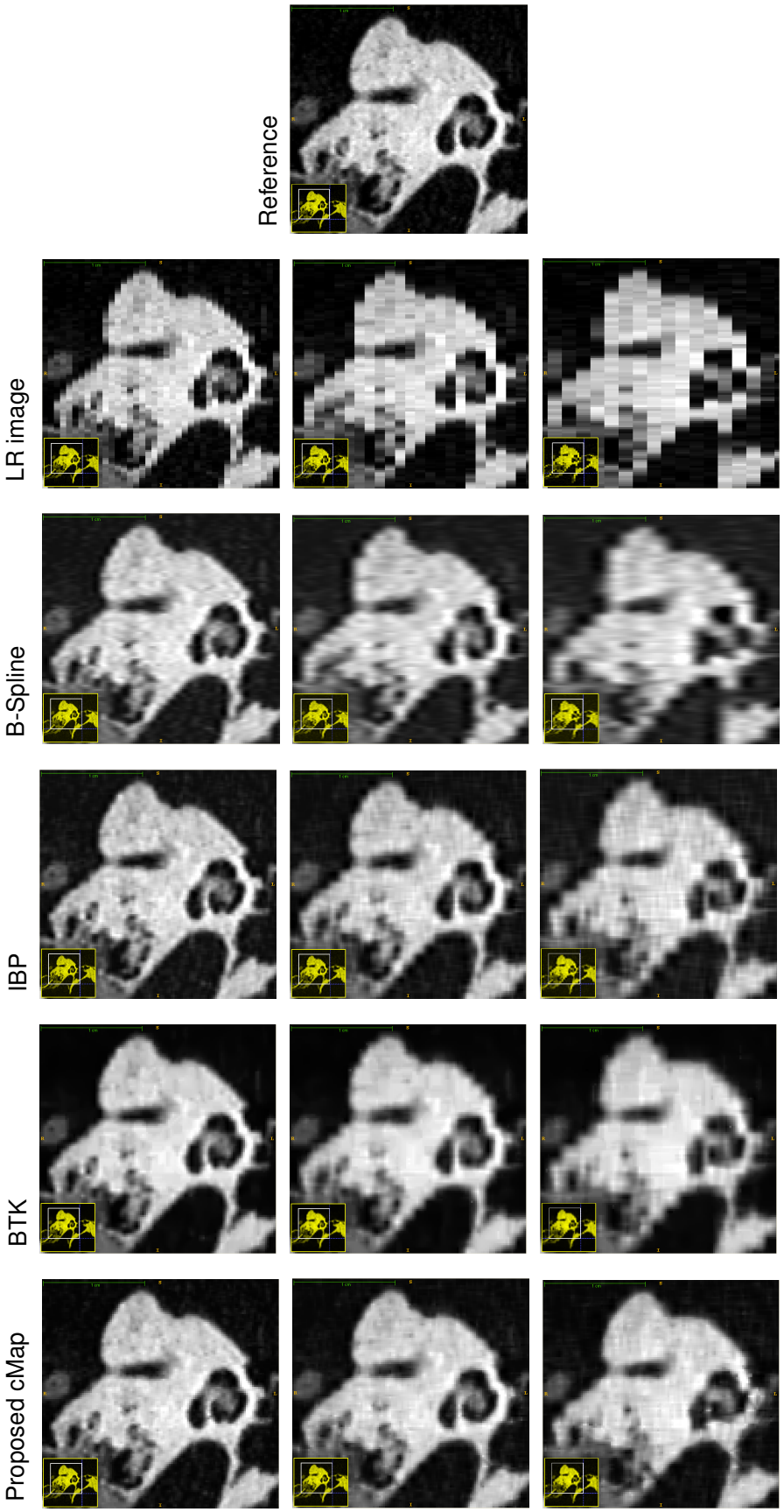
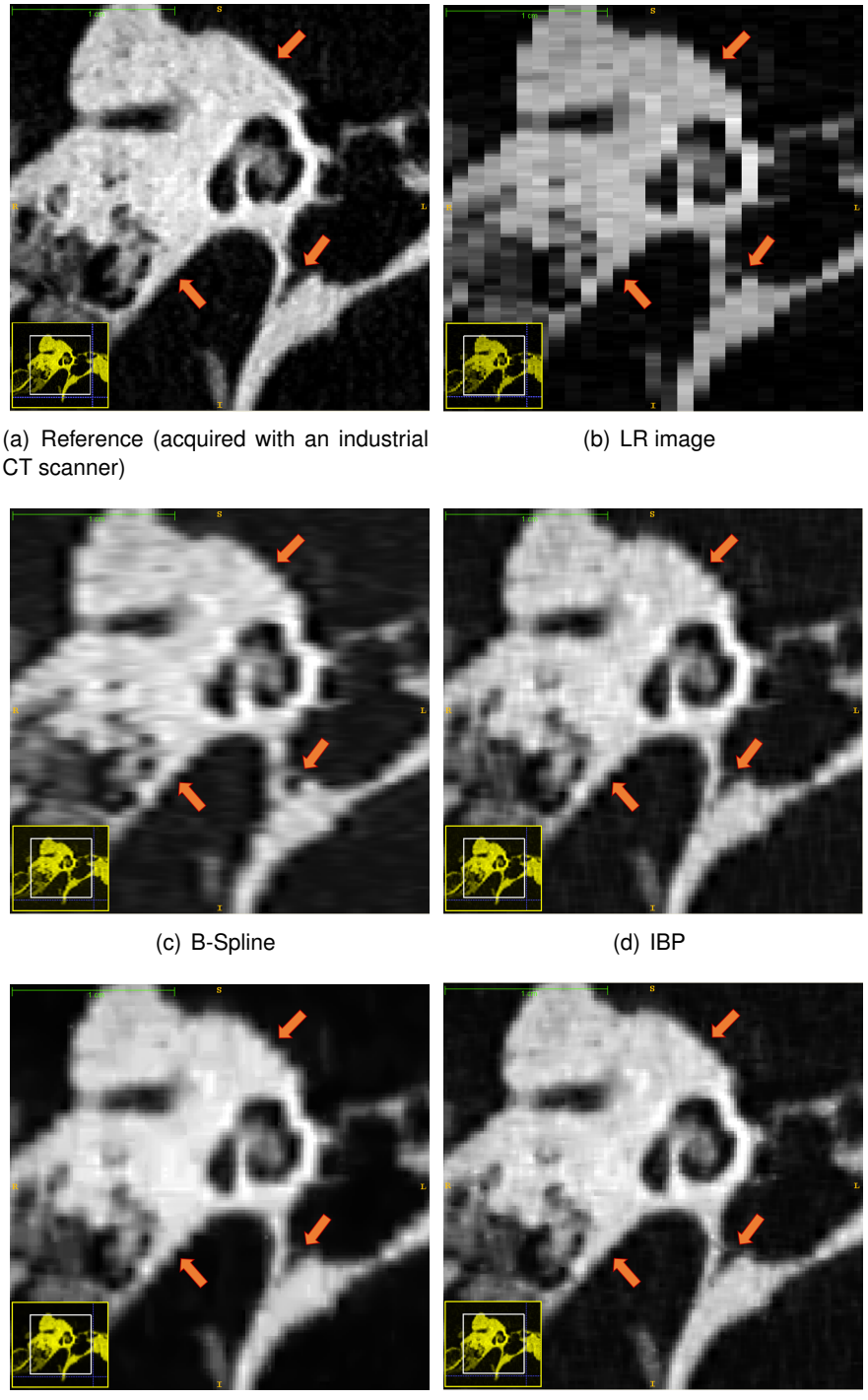


Figure 2.28.: Comparing reconstruction quality after applying various SRR techniques. The downsampling factors are 3 (left), 5 (middle), and 7 (right). The reference image was acquired with an industrial CT scanner



(a) Reference (acquired with an industrial CT scanner)

(b) LR image

(c) B-Spline

(d) IBP

(e) BTK

(f) Proposed cMap

Figure 2.29.: Comparing reconstruction quality after applying various SRR techniques. The out-of-plane resolution is five times lower than the in-plane resolution. A representative slice of the HR image and the appropriate LR image of one input data set are shown in the left column.

Robustness against noise. The performance of the proposed algorithm was evaluated in presence of noise using LR images corrupted with different noise models and levels. The two main noise types characteristic for CT images (Gaussian and Poisson noise) were considered. In these tests, the slice thickness of the input LR images was set to be five times larger than the in-plane spacing. The performance of the different reconstruction methods were quantitatively compared in term of PSNR. The corresponding results are plotted in Figures 2.30 and 2.31.

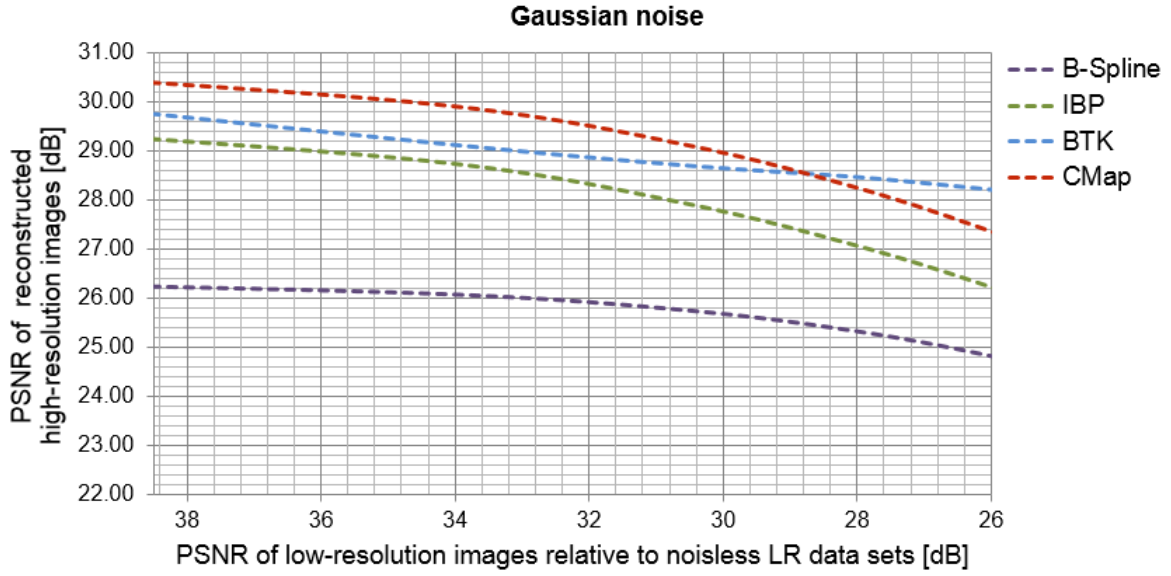


Figure 2.30.: Error profiles of the reconstructed images from LR data sets corrupted with Gaussian noise.

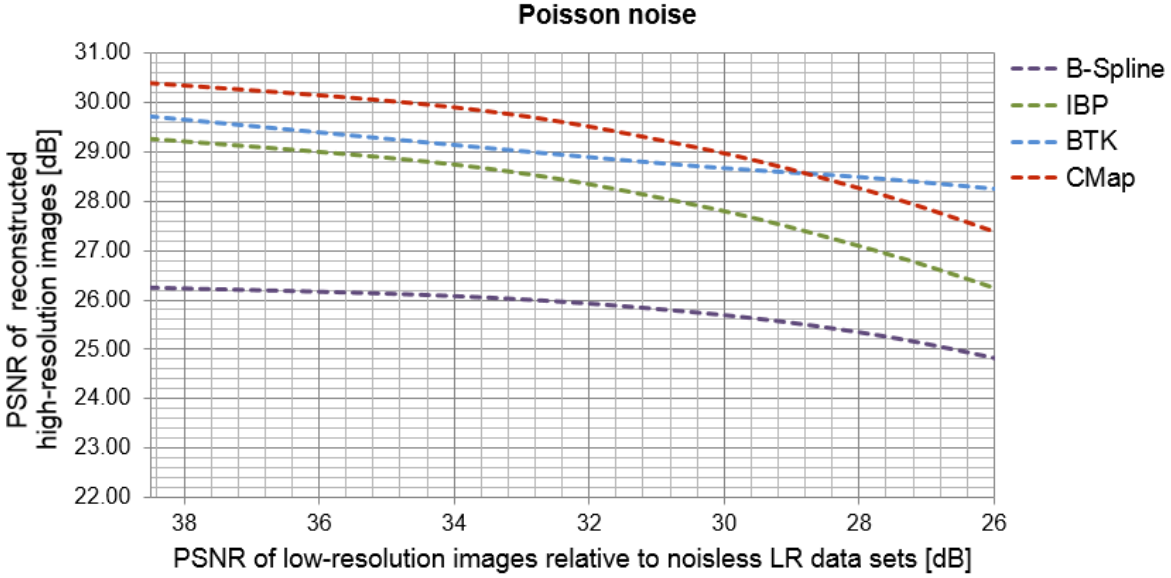
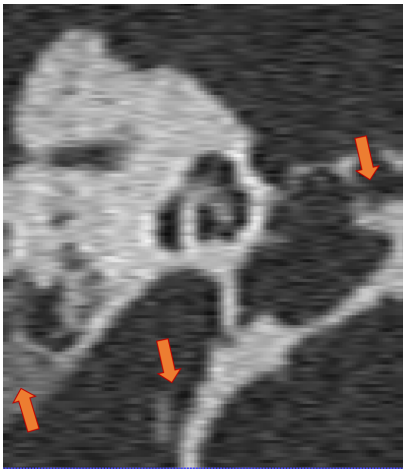


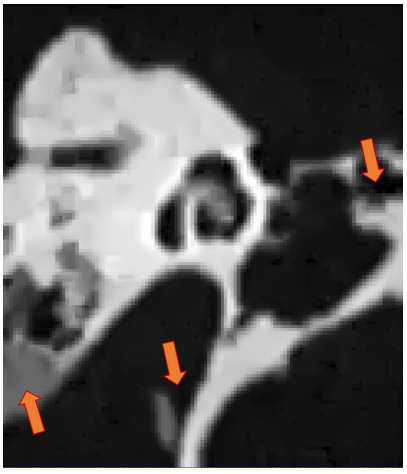
Figure 2.31.: Error profiles of the reconstructed images from LR data sets corrupted with Poisson noise.

(a) Reference (acquired with an industrial CT scanner)



(b) B-Spline

(c) IBP



(d) BTK

(e) Proposed cMap

Figure 2.32.: Comparing reconstruction quality after applying various SRR techniques. The low-resolution images are contaminated with high Poisson noise level (PSNR=24.47 dB). The anisotropy factor is set to five.

The error profiles computed in case of Gaussian and Poisson noise are very similar. It can be seen that for moderate noise level (> 29dB), the proposed method outperforms the other SRR approaches. In contrast, in case of excessive noise level (< 29dB) the M-estimation based SRR (BTK) has the best PSNR. However, the reconstruction result is not necessary visually better compared to the proposed approach. As illustrated in Figure 2.32, the M-estimation based SRR tends to over-smooth the HR volumes and results in a synthetic appearance of the reconstructed images. In contrast, the result achieved by the proposed cMap approach shows in different regions more details despite its noisy appearance.

It can be concluded that the noise present in the LR data sets has no major impact on the reconstructed HR images if its level remains in acceptable range, which is usually ensured by current scanners. For a moderate noise amount (> 30 dB) the proposed approach has always the lowest PSNR. In case of high noise level it is desirable to pre-process the input LR images in order to minimize the noise level.

Influence on the segmentation accuracy. In order to evaluate the influence of the reconstruction quality on the segmentation accuracy, step-cylinder phantom and double-ball-bar phantoms (see Figure 2.24) have been scanned with an isotropic resolution using an industrial CT. The scanned objects were segmented with the active contours method (also called snakes) using ITK-SNAP [PJH*06]. The segmentation parameters and the start points were set identical for all segmentation attempts. The segmentation error is defined as the deviation (Euclidean distance) between the segmentation performed based on the reconstructed SR images and the corresponding segmentation based on the acquired HR image. Therefor the surfaces of segmented regions were extracted. The mean error, root mean square (RMS), and the maximum Euclidean distance between the corresponding segmented surfaces are computed for each reconstructed SR volume. The computed errors were then scaled relative to the in-plane spacing, i.e. the desired high-resolution. As depicted in Figure 2.33, the best performance is achieved by the proposed cMap SRR. In addition, if the anisotropy factor is ≤ 5, which is ensured by modern spiral-CT scanners, the maximal segmentation error remains in the range of the in-plane spacing.

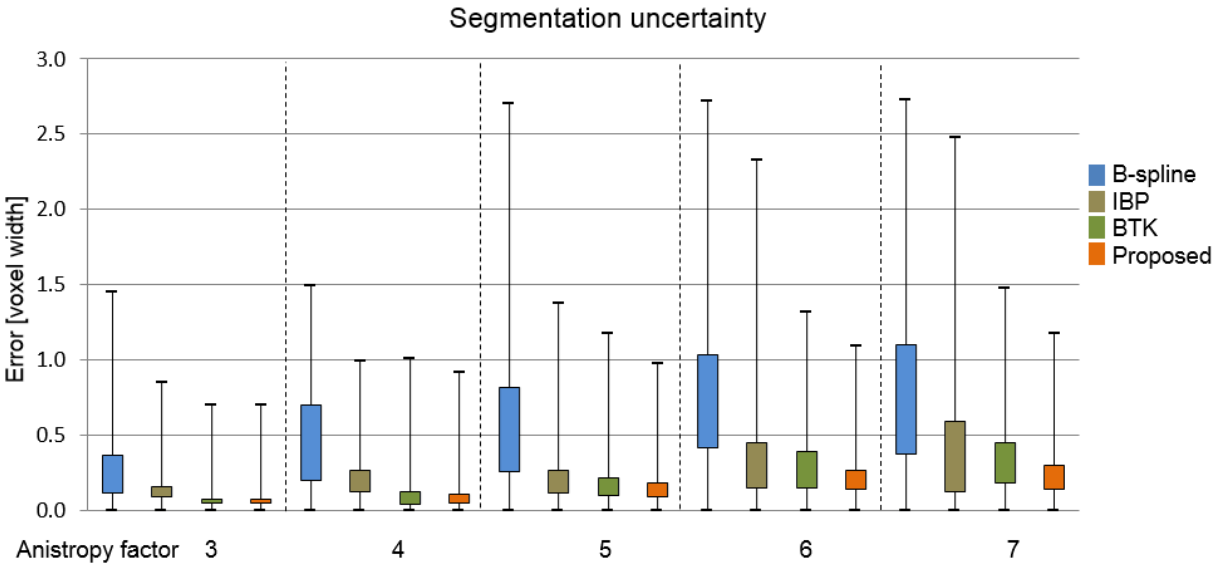


Figure 2.33.: Uncertainty of segmentation of PTFE spheres (placed in the region of interest) performed based on CTs reconstructed with different SRR methods and over different anisotropy factors.

2.5.2.2. Real data

Two nearly orthogonal CT scans of petrosal bone were acquired with a SOMATOM Sensation Open scanner (SIEMENS). The acquired low-resolution volumes have an in-plane resolution of 0.44×0.44 mm^2 and a slice thickness equal to 2.0 mm. As ground truth a third volume is acquired with an isotropic resolution using an industrial CT scanner. A comparison in term of PSNR is not possible since the image intensities are not calibrated with respect to Hounsfield scale in case of the industrial CT scanner. Therefore, the comparison was limited to visual inspection. Figures 2.35 and 2.36 show two representative slices of the reconstructed images. For comparison purposes, the corresponding HR slice and an original LR slice are also shown. It can be noticed that boundaries and fine details are better recovered if the proposed SRR approach (cMap) is applied.

Influence on the segmentation accuracy. In order to evaluate the possible improvement in the segmentation accuracy, two spheres made of PTFE are placed in the middle of the region of interest. The spheres are segmented in all reconstructed images and compared to the reference form. As shown in Figure 2.34, only the proposed SRR method could achieve the desired sub-pixel accuracy.

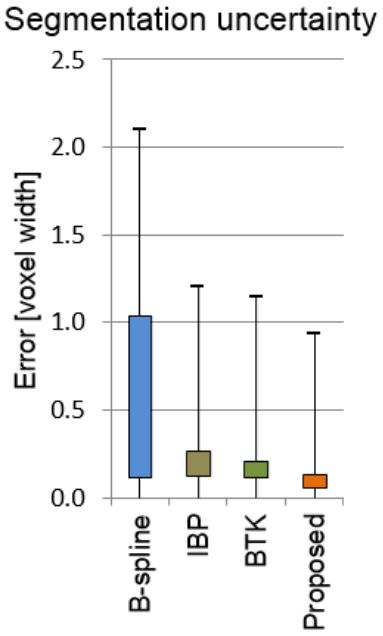
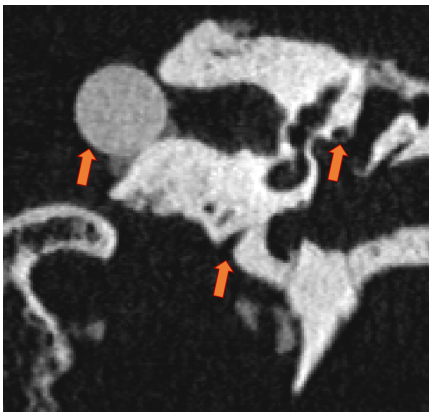


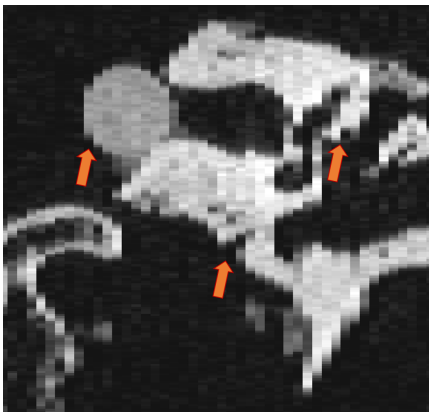
Figure 2.34.: Uncertainty of segmentation of PTFE spheres (placed in the region of interest) performed based on 3D images reconstructed with different SRR methods.

2.5.3. Computational complexity

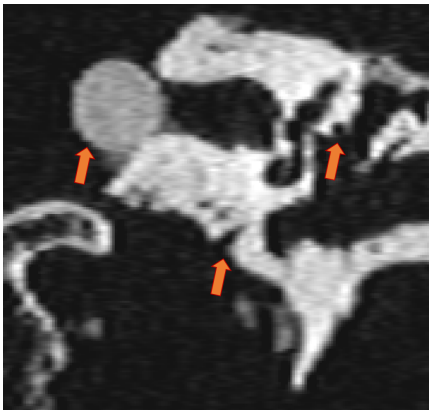
The experiments were performed on a system with an Intel Core i5-4670 CPU with 3,40 GHz CPU and 16 GBytes RAM. The mean processing time for a typical CT image size (512^3) is 13.39 seconds in case of IBP, 37.88 seconds by employing the proposed cMap, and 356.41 seconds by applying BTK. Thereby only the reconstruction process –without registration step– is considered. The classical IBP is the simplest approach and is therefore less time-consuming. The proposed approach requires more time but it still be acceptable for clinical application. Most of the running time is thereby spent for gaps inpainting. The most time consuming method is the BTK since it is computationally very demanding.



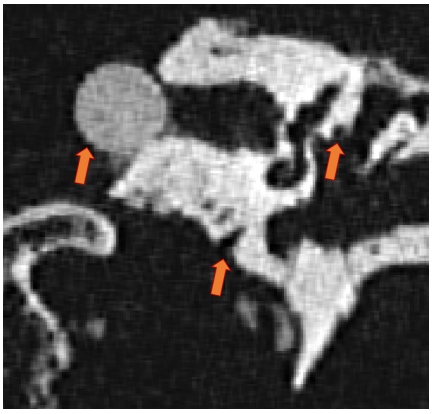
(a) Reference (acquired with an industrial CT scanner)



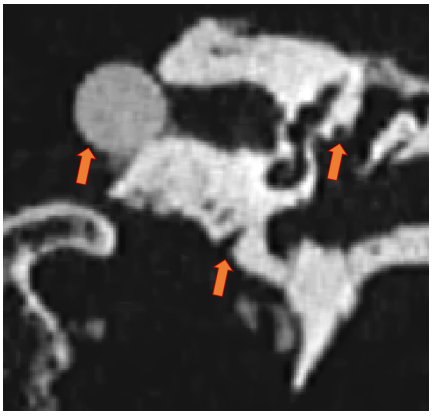
(b) LR image (acquired with a medical spiral CT scanner)



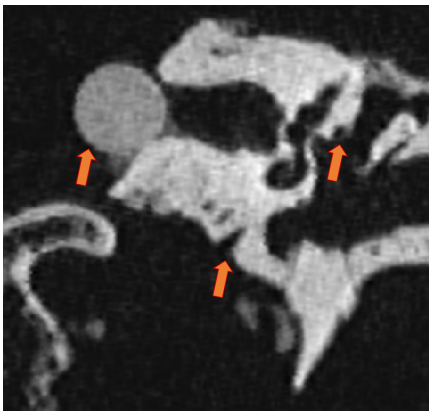
(c) B-Spline



(d) IBP

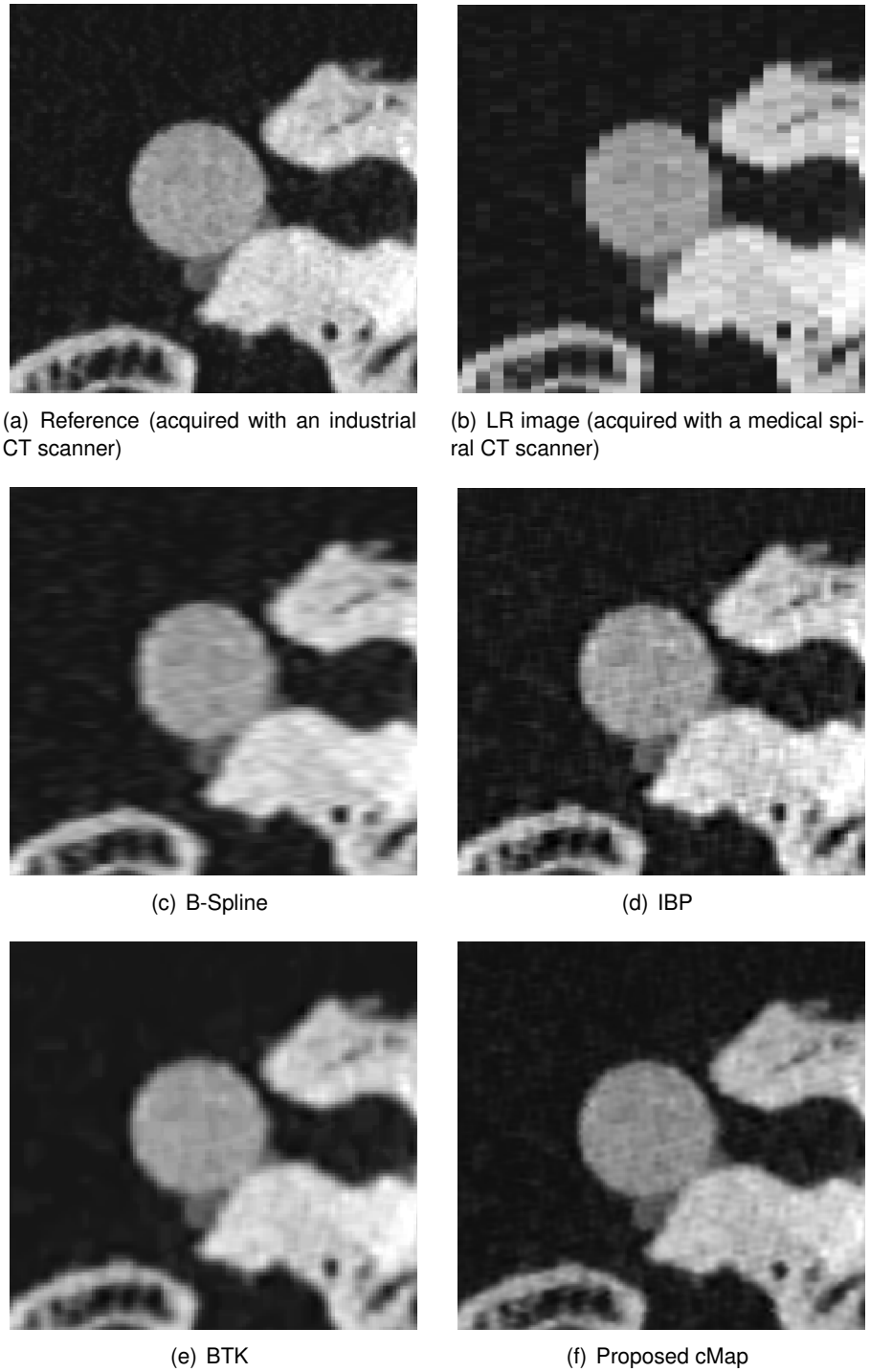


(e) BTK



(f) Proposed cMap

Figure 2.35.: Comparing reconstruction quality of various SRR methods applied to two nearly orthogonal low-resolution CT images (anisotropy factor $\approx$4.5). A representative slice of the HR image and the appropriate LR image of one input data set are shown in the left column.



(a) Reference (acquired with an industrial CT scanner)

(b) LR image (acquired with a medical spiral CT scanner)

(c) B-Spline

(d) IBP

(e) BTK

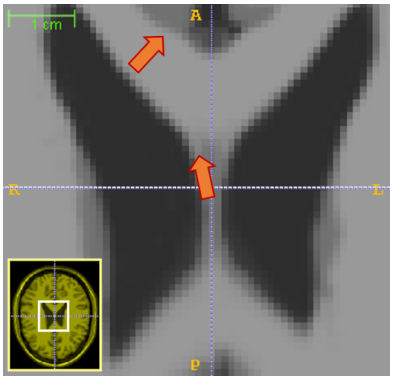
(f) Proposed cMap

Figure 2.36.: Comparing reconstruction quality of various SRR methods applied to two nearly orthogonal low-resolution CT images (anisotropy factor $\approx$4.5). A representative slice of the HR image and the appropriate LR image of one input data set are shown in the left column.

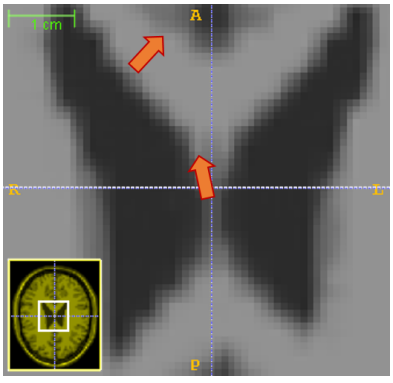
2.5.4. Magnetic resonance imaging

Acquiring MR images with high resolution is always achieved at the expense of the acquisition time or the signal-to-noise ratio. Thus, it is common to acquire a set of 2D slices, not a true 3D image. Super-resolution reconstruction can therefore be applied to reconstruct 3D images with isotropic resolution from the acquired LR data sets. The adequacy of the proposed SRR approach for Magnetic resonance imaging was also tested. The MR images are obtained from the BrainWeb database [CZK*98].

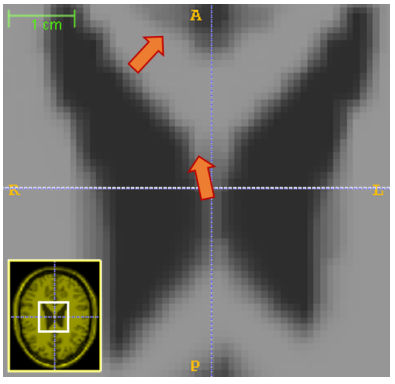
It is to be noted that MR images may suffer from intensity inhomogeneity leading to reconstruction artifacts if SRR techniques are applied to the original images. Thus, it is mandatory to apply a histogram matching [WLB*98] in order to correct the inter-intensity inhomogeneity. Therefore, the first data set is considered as reference image and then the intensities of the second image are redistributed to match with the intensities of the first data set. The corrected images are then used as input for the investigated SRR methods. Figure 2.38 shows slices of axial, coronal and sagittal views. A zoom on a small detail is shown in Figure 2.37. It can be seen that sharper details with enhanced contrast can be obtained with the proposed SRR approach.



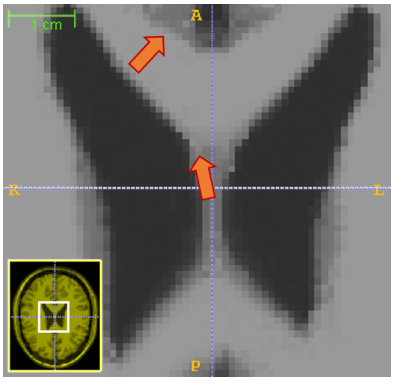
(a) Reference



(b) IBP

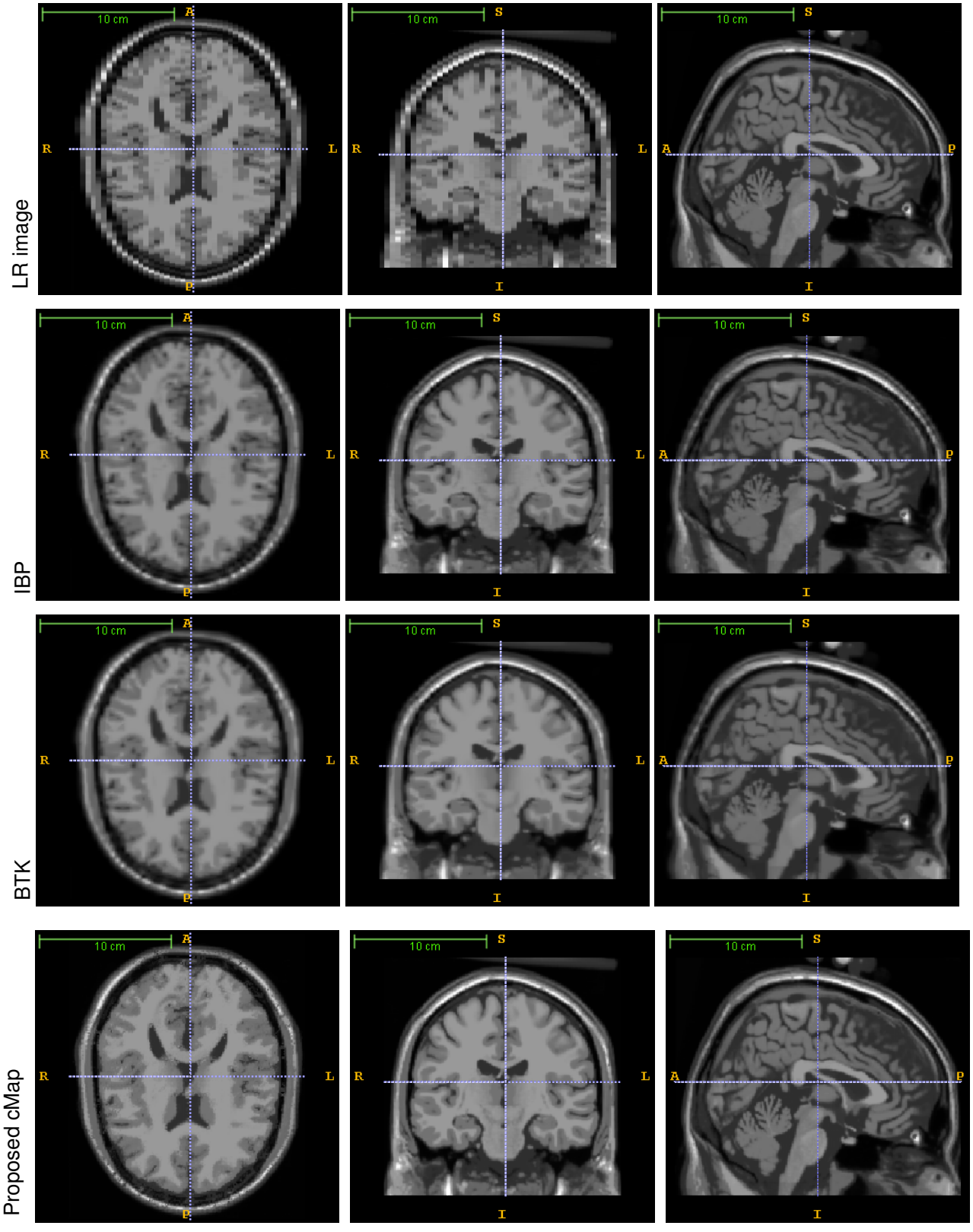


(c) BTK



(d) Proposed cMap

Figure 2.37.: Comparing reconstruction quality of small details after applying various SRR techniques on MR images.



(a) axial (b) coronal (c) sagittal

Figure 2.38.: Axial, coronal, and sagittal views of one LR MR image and the appropriate views of reconstructed images.

2.6. Summary and discussion

In this chapter an algorithm to reconstruct an isotropic high-resolution image from two orthogonal low-resolution data sets was presented. The application of the proposed confidence-map based SRR in CT as well in MRI has been investigated. The focus was set on the reconstruction of appropriate HR images for lateral-skull minimal invasive surgery.

The new approach aims at estimating the confidence of the input images and then recovering the HR image by maximizing the data fidelity to the high-confident information. A main contribution of this work is the estimation of the uncertainty of the input information, which is considered in the reconstruction process. An improved initial estimation of the unknown HR image serves for faster convergence against a better regularized solution.

By applying the proposed method, fine details of the anatomical structures could be accurately reconstructed. Consequently, the segmentation accuracy would be improved. The peak signal-to-noise ratio is usually used as a quantitative measure for the reconstruction quality. However, this measure is not appropriate to validate the resolution enhancement or the reconstruction accuracy. Therefore, we proposed to study the influence of the reconstruction improvement on the segmentation uncertainty since this later is crucial for the planing and navigation stuff. It was shown that the maximal deviation in all direction could be kept smaller than the in-plane spacing by applying the proposed SRR method if the anisotropy factor is less or equal to five.

In addition, the presented approach has the advantage of allowing the reconstruction of HR images from already reconstructed 3D CT images. This offers more flexibility to fuse data sets, acquired even with different scanners or over different time points, as a post-processing step. The method relies on conventional CT images without requiring additional hardware. Furthermore, the SRR is performed fully automatically, so that after acquiring the low-resolution images –as described in Section 2.4– no further user interaction is needed to complete the reconstruction process. On the other hand, the proposed approach is computationally effective due to the enhanced initialization and the simple iterative update approach. All these aspects make the proposed SRR appropriate for clinical applications and easy to be included in clinical suites.

Outlook

In practice the number of LR images is limited due to time and radiation constrains. Therefore, the number of the input low-resolution data sets was fixed to two images. Nevertheless, it is possible to apply the proposed approach to any number of input images. Furthermore, the proposed confidence map can be easily extended to include further prior knowledge about the imaging system and reconstructed artifacts in the input data sets.

In order to enhance the segmentation accuracy, it may be beneficial to incorporate the estimated confidence-map in the segmentation process itself. The confidence map can also be used to provide a visual feedback about uncertain regions which should not be trusted during the diagnosis. This may help to avoid wrong diagnosis due to the low image resolution and image artifacts.

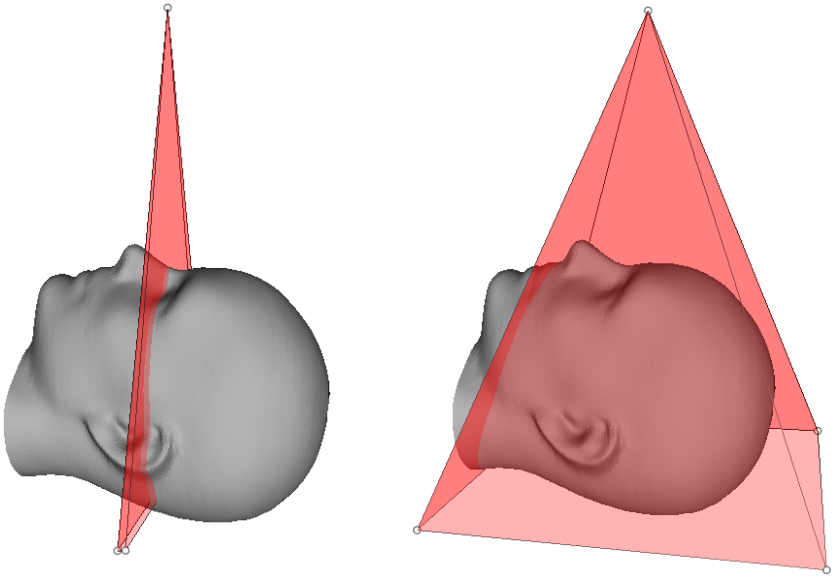
The blurring function is usually unknown and is therefore simplified. Developing an accurate point spread function (PSF) model for spiral-CT scanners may improve the reconstruction quality if it is incorporated in the SRR algorithm.

The performance of SRR depends on the accuracy of the registration process. In order to perform a robust and accurate alignment of the LR images fiducial markers are used. It would be desirable to develop a sub-voxel accurate markerless registration approach for anisotropic images.

3. CBCT reconstruction

3.1. Introduction

Spiral CT scanners are not appropriate for intra-operative data acquisition and are therefore usually limited to perform pre- or post-operative scans. In contrast, C-arm CBCT systems are designed to be suitable for interventional suites. Most recent C-arm scanners are equipped with flat-panel detectors (FPD) enabling the acquisition of 3D images based on X-ray images gathered in a single orbit about the patient. Thereby the fundamental physical concepts are similar to conventional fan-beam CT techniques. A main difference between both systems is the projection geometry. In cone-beam CT systems, the X-ray beam does not form a fan-beam geometry but rather a conical geometry as depicted in Figure 3.1.



(a) fan beam geometry (b) cone-beam geometry

Figure 3.1.: Illustration of fan beam geometry typical for spiral CT scanners (a) and cone-beam geometry characteristic for C-arm CBCT systems (b).

C-Arm CBCT systems can theoretically acquire images with high isotropic spatial resolution and relatively low radiation dose (compared to spiral CT systems). In practice, the reconstructed 3D images usually suffer from CT artifacts due to several limitations in the imaging system and reconstruction process. For instance, current scanners employ usually the well-known Feldkamp reconstruction algorithm [FDK84] due to its simplicity and low computational cost. However, this algorithm has several limitations:

1. it is not appropriate for large cone angles since the image quality would be degraded for regions outside the central plane [KK07]
2. substantial quality degradation and heavy reconstruction artifacts are expected if metal implants are in the scanned region
3. ideal circular scan trajectory is required. This assumption is usually not satisfied in case of C-arm systems due to large mechanical inaccuracies.

This chapter is structured as following: Section 3.2 presents the physical principles underlying the CBCT technology and describes common reconstruction artifacts. Section 3.3 reviews the state of the art of CBCT reconstruction algorithms. Section 3.4 introduces several new algorithms dealing with two main aspects in CBCT, namely the on-the-fly geometrical calibration and the reduction of metal-related artifacts in the reconstructed 3D CT images. Section 3.8 discusses the results of the proposed algorithms and proposes recommends for future work.

Notation. The basic mathematical conventions used in this thesis are listed below.

- Vectors and matrices are denoted by bold capital letters, e.g. $\mathbf{A}$. An element on the ith row and jth column of a matrix $\mathbf{A}$ is denoted by a_{ij}
- Direction vectors are indicated with a right-handed arrow over the letter, e.g. $\vec{r}q$ to describe the projection direction of the X-ray captured by the pixel q
- Volumes and images are denoted by capital letters, e.g. V for volumes and P for X-ray images.
- Volumes as well as images are lexicographically ordered, which results in (1D) vector containing all voxels or pixels in one single row, respectively. Voxels and pixels can therefore be addressed as $V(x, y, z)$ or $V(i)$. Similarly, pixels can be addressed as $P(u, v)$ or $P(q)$
- A particular image of the X-ray images P acquired at the gantry angle θ is written P_θ
- Simulated data are indicated by tildes over letters, e.g $\widetilde{P}$ for simulated projection images (DRRs)
- Mean values are indicated by bars over letters, e.g $\bar{P}$ for the mean values of all X-ray pixels
- The superscript $^{(k)}$ is used for denoting the updated estimate e.g of images at the kth iteration, like $V^{(k)}$

3.2. Background

3.2.1. Mobile C-arm CBCT system

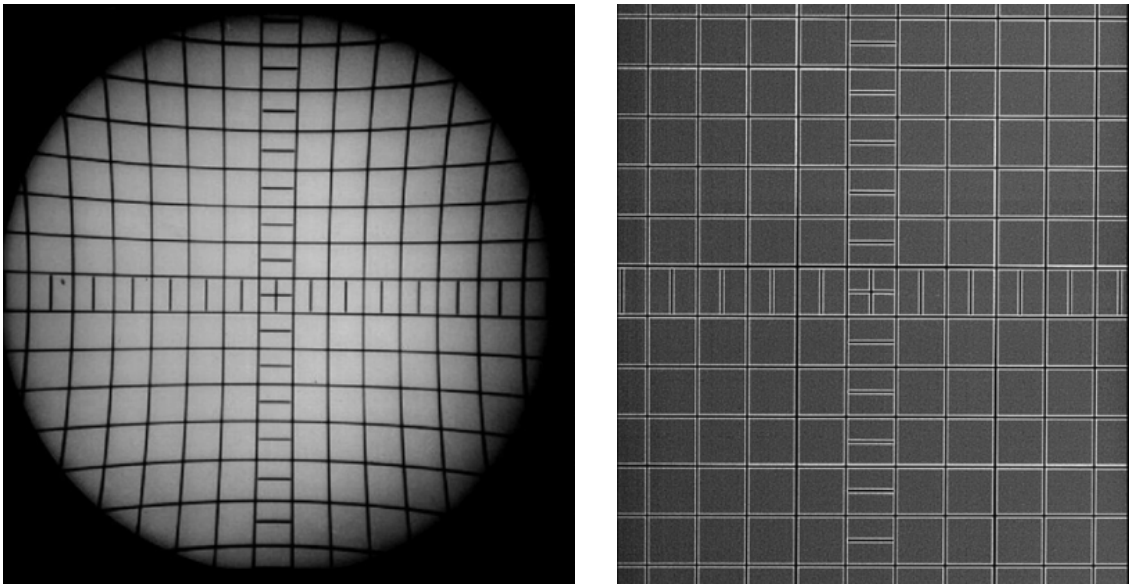
A special prototype of C-Arm CBCT system was provided by Ziehm Imaging GmbH (Nürnberg, Germany). It consists of an X-ray source-detector combination as depicted in Figure 3.2. The X-ray source and the flat panel detector (FPD) are mounted in the opposing ends of the open arm and have a fixed distance. The nominal source-to-detector distance (SDD) is 106.4 cm. The flat panel detector size is 30 cm $\times$ 30 cm and has a resolution of 0.283 mm $\times$ 0.283 mm (1024 $\times$ 1024 pixels).

The X-ray images are acquired using a single gantry with a limited short scan angular range of 165°. For 3D imaging, an automated orbital movement can be performed around a variable iso-center and follows thereby a circular or an elliptical trajectory. The device is equipped with a 7.5 kW generator which can sustain a pulsed radiation from 40 to 120 kV tube voltage with a tube current from 0.1 to 75 mA and a pulse

Figure 3.2.: C-arm CBCT system.

width from 4 to 50 ms. In addition, an Object Detected Dose Control (ODDC) serves for an automatic dose reduction and contrast correction.

3.2.2. Flat panel detector



(a) Image intensifiers (IIT)

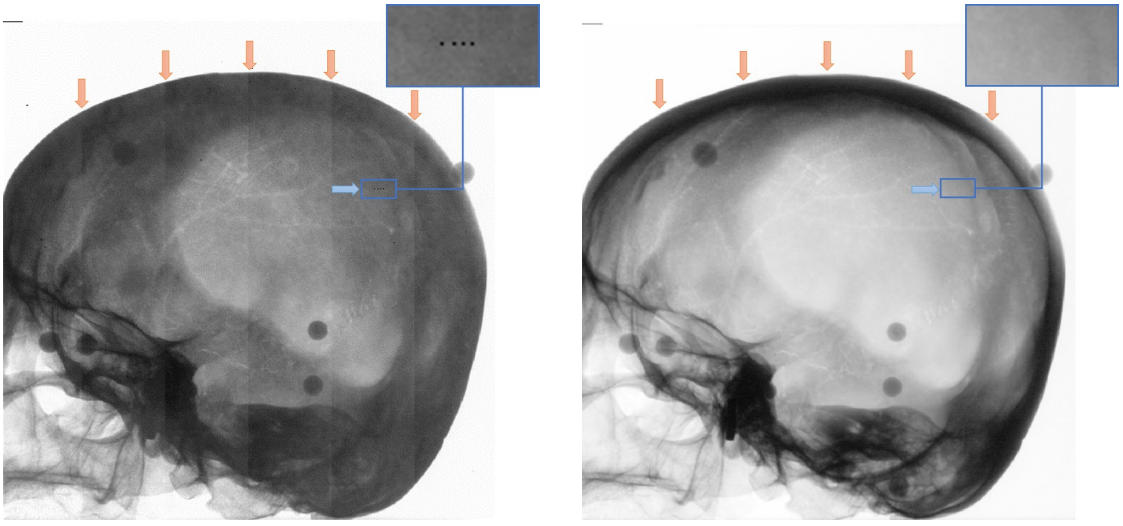
(b) Flat panel detector (FPD)

Figure 3.3.: Comparison between 2D X-ray images acquired with an image intensifier (a) and a flat panel detector (b). Unlike the image intensifier scan, the scan acquired with the flat panel detector is not affected by distortion and brightness non-uniformity [CDS08].

Flat panel detectors (FPD) follow an indirect measurement technique: the X-ray photons hit a layer of scintillating material (CsI) and are converted into visible light photons. These photons strike then photodiodes of amorphous silicon which generate electrical signal in proportion to the captured light [KR02].

An obvious advantage of flat panel detectors compared to image intensifiers (IIT) is the decrease in size and weight. Thus, it is possible to mount a FPD with a larger active area on a mobile C-arm system. In addition, unlike image intensifiers, FPDs are able to acquire distortion-free images. They have a direct short signal conversion path, which avoid possible distortions caused by the earth's magnetic field, such as the geometric distortion and the brightness non-uniformity. An example illustrating the image distortion is shown in Figure 3.3. A further advantage of FPDs is their almost linear response signal covering bigger dynamic range in comparison to IITs.

The signal produced by the absorbed X-rays should be higher than the background noise but lower than the detector saturation point. Therefore, a signal amplification is performed prior to the digitizing process to achieve a reasonable signal level for digitization. Further corrections, such as offset and temperature correction, have to be applied to the raw data in order to compensate for non-uniformities in the detector and generate usable 2D X-ray images. An example illustrating the enhancement in quality by performing detector calibration is shown in Figure 3.4. The main corrections methods are reviewed and discussed in [Chr08].



(a) Original acquired CBCT image (b) Corrected CBCT image

Figure 3.4.: An example illustrating the enhancement in image quality through defect-cells, detector offset and temperature correction. A zoom (blue arrow) illustrates the correction of defect pixels. Red arrows point out the inhomogeneity correction.

3.2.3. Technical limitations

C-arm CBCT systems have some significant technical limitations resulting in quality degradation in the reconstructed 3D images. In the following a summary of the main limitations is presented.

1) Mechanical inaccuracy. Mobile C-Arm systems are mechanically unstable in comparison to spiral CT scanners. They exhibit large geometric deviations (up to several millimeters) due to mechanical flex. Such mechanical inaccuracies lead to deviation from the expected ideal scan orbit (Figure 3.17) and hinder therefore precise 3D reconstructions. Under the assumption that the geometrical deviations are repro-

ducible, an appropriate offline geometric calibration can be performed to compensate the mentioned error. Actually a short term reproducibility in sub-millimeter range could be ensured [DSC*08]. However, this is still insufficient to achieve high accurate reconstructions since the composed geometrical error in the projection area is larger than the detector pixel width. Furthermore, an offline calibration does not account for stochastic-mechanical inaccuracies (vibrations) occurring during the scan process.

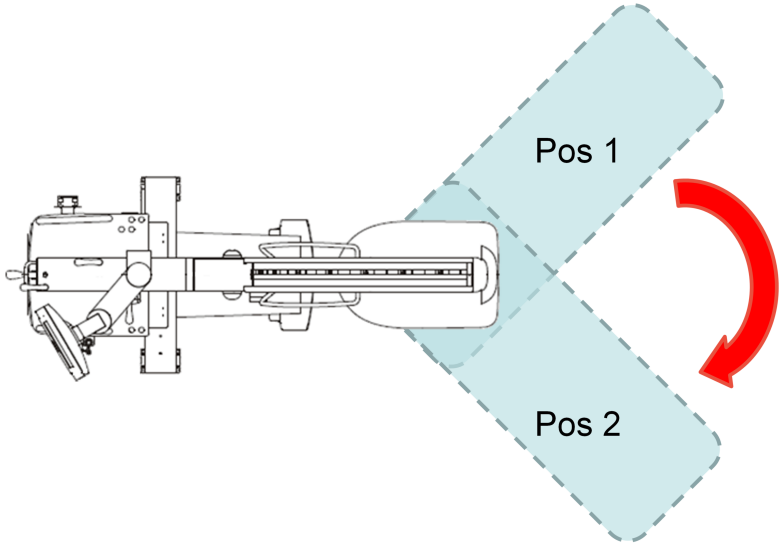
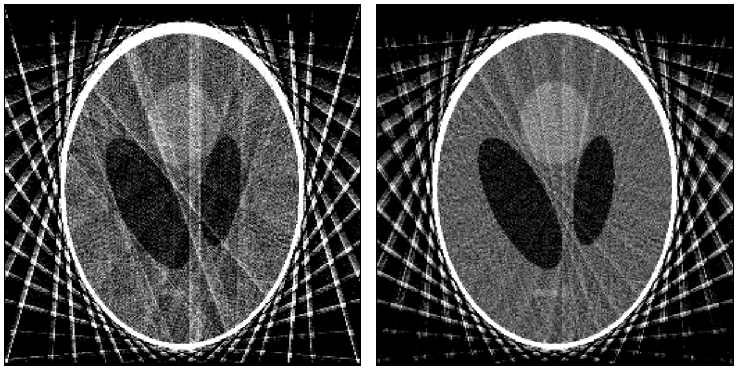


Figure 3.5.: Acquisition of two complementary data sets in order to compensate the missing projection angles.

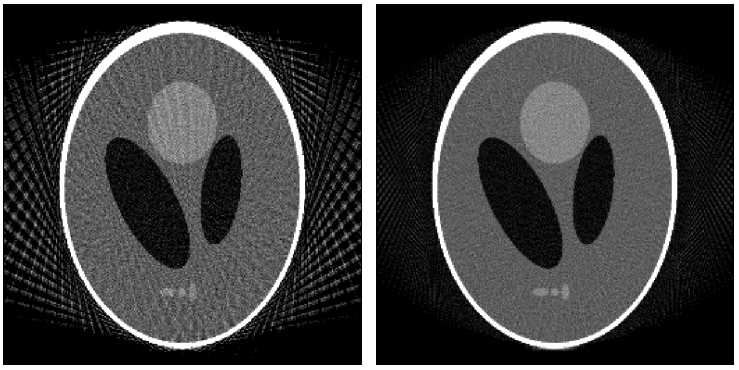
2) Gantry angle. Theoretically the required gantry angle should be over 180°plus fan angle. This requirement is not satisfied in most of C-arm devices. According to Tuy's data sufficiency condition [Tuy83], the data acquired following a circular scan trajectory is incomplete for an exact reconstruction in the areas outside the trajectory-plane. For instance, the planes parallel to the trajectory-plane are not intersected by scan orbit. Thus, only an approximate reconstruction can be performed.

The orbital range of the used C-arm CBCT system is 165°, which is less than the 180°plus fan angle required for accurate tomographic reconstruction. In order to overcome this problem and reduce the expected image artifacts due to limited angle orbit, it is possible to combine two scans along two orthogonal orbits as depicted in Figure 3.5.

3) Number of acquired X-ray images. Another limitation is the relatively low number of acquired X-ray Images. Ideally, the number of gathered images should be adjusted so that the angular separation between two successive projections amounts to a distance at the border of the field of view (*FOV*) equal to the voxel size (V_{size}). Mathematically, this means that the number of acquired X-ray projections should be equal or larger than $\frac{2\pi}{\arctan(2V_{Size}/FOV)}$ [JSW81, ZGS14]. For example, to reconstruct a CT volume which consists of 512^3 voxels, ca. 3200 X-ray projections are required. Acquiring such number of projections is usually not feasible due to technical and time constrains. Furthermore, the radiation dose must be kept as low as possible. For illustration purposes, a Shepp-Logan phantom was reconstructed based on different projection numbers as depicted in Figure 3.6.



(a) 18 (b) 36



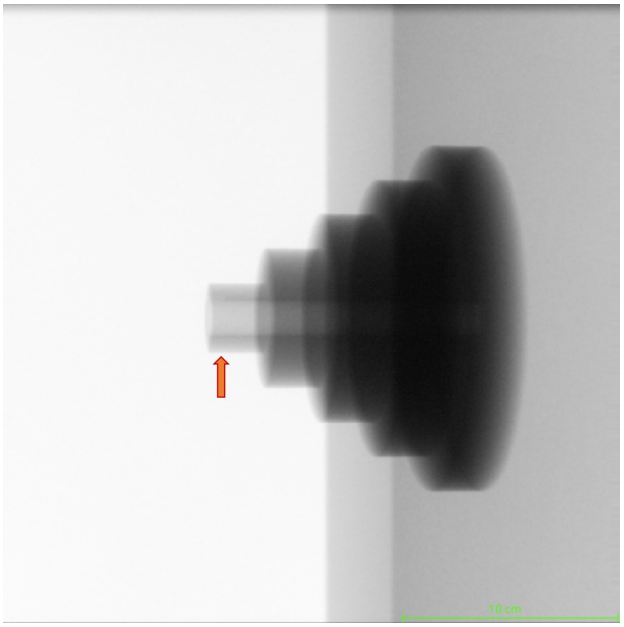
(c) 72 (d) 180

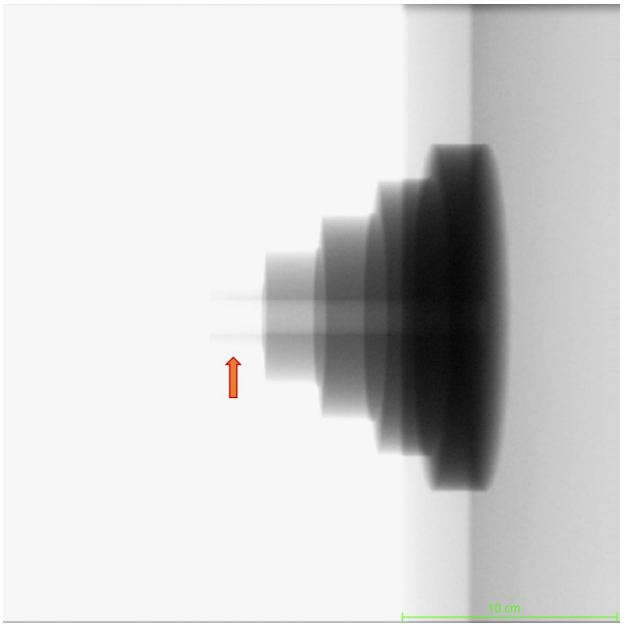
Figure 3.6.: An example illustrating the influence of the number of the projection images on the reconstruction quality. Representative slices of 3D reconstructions of a Shepp-Logan phantom based on 18 (a), 36 (b), 72 (c) and 180 (d) discrete angles along a 360° arc around the scanned phantom are shown.

4) Scatter. Scattered photons can hit the detector at positions different than the direct beam. This effect happens when photons are diffracted from the direct path after interaction with the scanned object. Due to the relatively large area of used flat panel detectors, the probability to measure scattered photons is higher than in conventional spiral-CTs. Effectively, the scatter-to-primary ratio (SPR) in case of large FPDs exceeds 2.5 for big volumes [SZB*13]. For comparison, multi-detector CTs have SPRs of around 0.2. It was also experimentally shown that the SPR increases significantly while increasing the cone angle [SJ01].

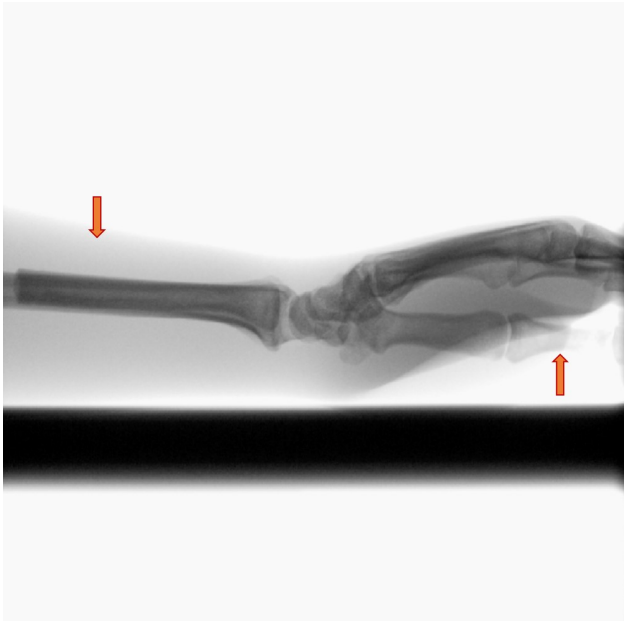
The increased scatter level causes image artifacts such as streaks. In order to minimize the effect of scattered radiation on the acquired X-ray images, anti-scatter grids are employed. However, this may cause a decrease in contrast-to-noise ratio (CNR) [SMB*04]. Some scatter correction algorithms were proposed to deal with this scatter effect [NTC04] [SZB*13].

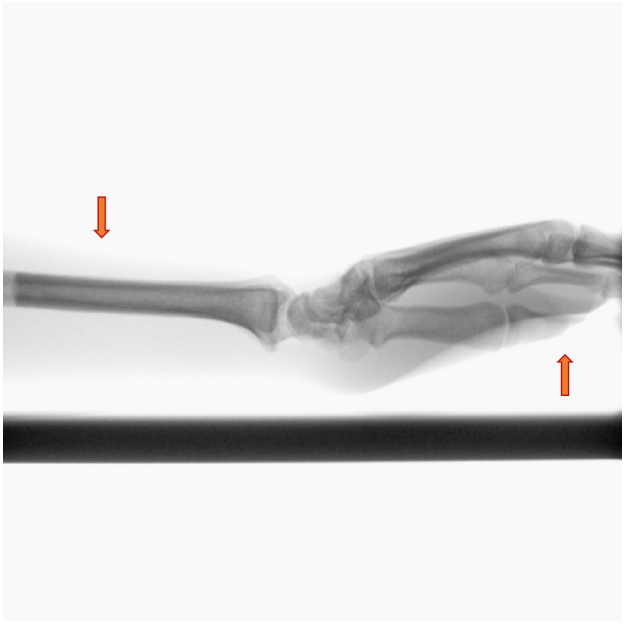
5) Noise. The acquired data is usually corrupted by electronic noise and photon-count noise which are expected to follow a Poisson distribution [YLL*09]. This problem is more evident in CBCT systems since they operate at milliampere settings, i.e. one order of magnitude below conventional spiral-CT scanners. CBCT machines have therefore much lower signal to noise ratios, which typically causes a degradation in reconstruction quality.





(a) Two successive 2D X-ray images of a step-cylinder-phantom. The upper cylinder part is not visible in the second projection due to the over-estimation of needed X-ray radiation dose.





(b) Two successive 2D X-ray images of a hand-phantom. The contrast in the soft tissue area is lower in the right projection than in the left image.

Figure 3.7.: Illustration of image quality degradation caused by over-estimation of needed X-ray radiation dose.

6) Inhomogeneity. The quality of the X-ray images may suffer under the limited dynamic range of the flat panel detector. Thus, an automatic dose regulation is usually employed to continuously adapt the X-ray radiation to the thickness and density of the irradiated object. However, this technique may result in additional quality degradation in case of strong variation in the attenuation integrals between successive projections. In this case, the dose can be over- (saturation) or under-estimated (under-exposure) due to

abrupt variations in the required dose, which may result in bright or dark areas with unusable data in some regions of the acquired X-ray projections. This problem is illustrated through two examples shown in Figure 3.7.

3.2.4. Reconstruction artifacts

In this context, reconstruction artifacts are defined as reconstructed details which are not present in the scanned objects. They appear often as streaks and shadows oriented along the projection rays. Such artifacts are basically caused by discrepancies between the modeling and the acquisition process. In this section, a brief summary of common artifacts and their causing factors is presented. A detailed review can be found in [KHE87, SHG*11].

1) Ring artifact. Ring artifacts appears as concentric rings around the rotation axis. They are basically caused by uncalibrated or defective detector elements. An example illustrating such artifacts is shown in Figure 3.8.

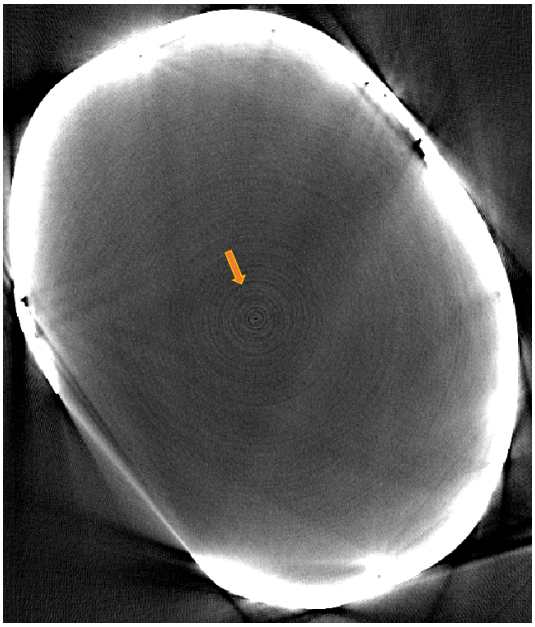


Figure 3.8.: Ring artifacts centered around the rotation axis.

Figure 3.9.: Example of beam hardening artifacts in CT images.

2) Beam hardening. In practice, the X-ray source emits polychromatic spectrum. However, the reconstruction algorithms are usually simplified by assuming the X-ray source to be monochromatic. Moreover, the attenuation of the X-rays depends on the energy spectrum. Effectively, low energy X-rays, also called soft X-rays, are more rapidly attenuated compared to high energy X-rays. The lower energetic rays of this spectrum are substantially absorbed when a high energetic part of the X-ray beam passes though the object under investigation, which results in increasing the average energy level of this beam since its lower-energy parts are more rapidly attenuated than the higher energy parts. This effect causes beam hardening artifacts which appear as dark streaks in the CT images as depicted in Figure 3.9.

Overcoming such artifacts can be achieved by one of the following mechanisms:

- Pre-filtering: physically pre-filter the X-ray beams in order to eliminate the lower energetic parts
- Modeling polychromatic X-rays: incorporate a polychromatic acquisition model in the 3D reconstruction process
- Post-processing: correct the reconstructed CT image based on material assumptions

3) Scatter. It is usually assumed that the projection values are exclusively related to non deflected X-rays, i.e. no scattered photons contribute to the recorded projection values. This assumption is violated in case of high energy X-rays, due to the high number of scattered photons with enough energy to penetrate through the scanned object and contribute to a detector pixel. This effect impacts basically the low-intensity measurements, i.e. the pixel values corresponding to the X-rays attenuated by highly absorbing materials (high-density objects). Scatter artifacts appear in form of streak artifacts and reduced soft-tissue contrast as illustrated in Figure 3.10.

4) Exponential edge gradient effect. Due to the finite focal spot width and the detector cell size, the acquired data are measured over finite beam widths. In the reconstruction process this width is assumed to be zero. This results in streaks which are tangent to sharp edges with high contrast to the surrounding structures. An illustration of such artifacts is shown in Figure 3.11.

Figure 3.10.: An Example of scatter artifacts with typical lowered and inhomogeneous CT values.

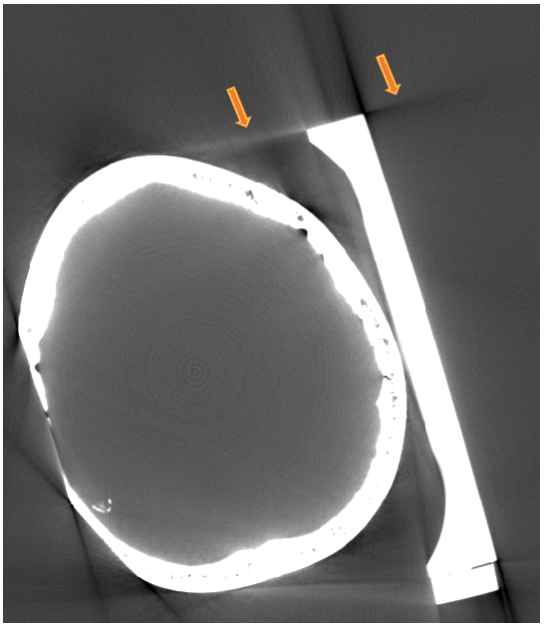


Figure 3.11.: Typical appearance of exponential edge-gradient effect characterized by thin lines tangent to sharp edges.

5) Aliasing. Considering the fundamental Nyquist sampling theorem, the sampling frequency must be higher than twice the highest signal frequency. Violating this requirement leads to aliasing artifacts in the reconstructed images. Therefore, the reconstruction process should take the detector pixel size and the divergence of the cone beam into account.

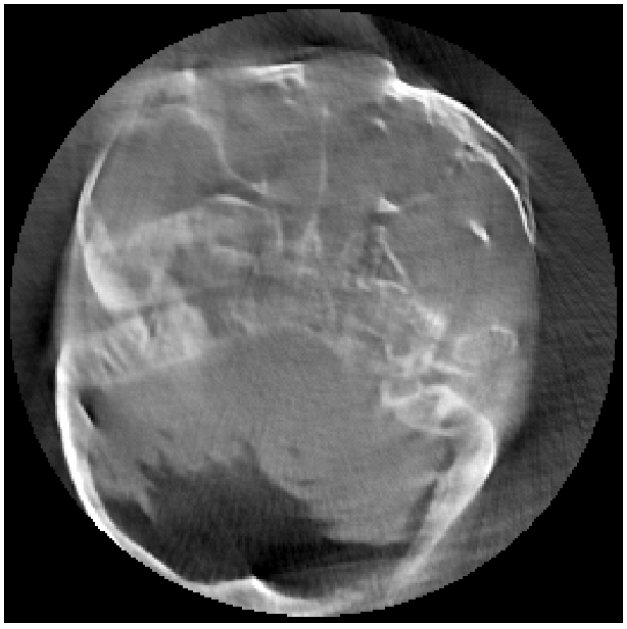


Figure 3.12.: Example of CT artifacts induced by patient movement during the data acquisition.

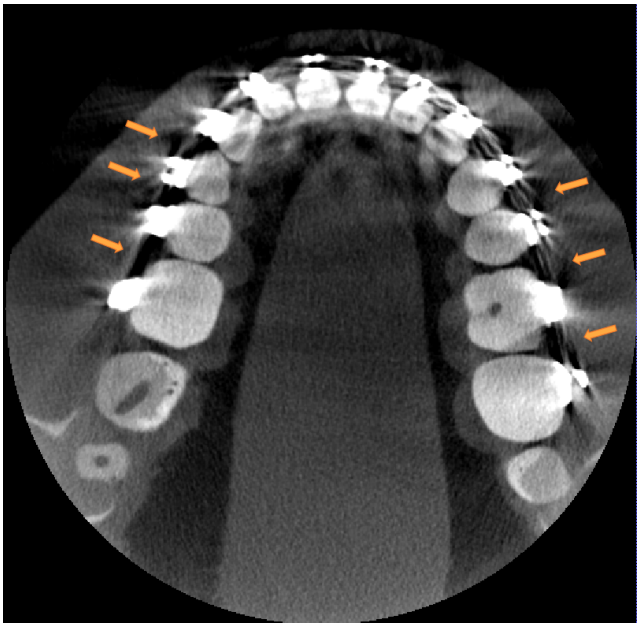


Figure 3.13.: Typical appearance of metal related artifacts in form of streaks an shadows along the projection rays.

6) Misalignment artifacts. Motion of sharp edges during the acquisition of the projection images, i.e. caused by patient motion, jittering or wrong iso-center, causes reconstruction artifacts if the misalignment of the components (source, detector, object) is not considered. This is evident since the back-projection is applied along lines which do not correspond to the recorded X-rays directions. Such artifacts appear typically as double contours as depicted in Figure 3.12.

7) Metal artifacts. The presence of metal objects in the scan field can lead to severe artifacts which appear as streaks an shadows. Beam hardening and scatter both produce dark streaks between high density objects, which look like balk shadows as shown in Figure 3.13.

Metal artifacts can be reduced by scanning at a higher kV, since harder X-rays are less likely to be completely attenuated by metal implants. However, this would result in less soft tissue contrast and cause additional radiation exposure for the patient. A better alternative is to address this problem using iterative reconstruction techniques as explained in Section 3.3.5.

3.3. Related Work

3.3.1. CBCT reconstruction

Given a set of X-ray images P_θ acquired at several gantry angles θ with q ($= u \times v$) pixels each, every projection pixel $P_\theta(q)$ is the line integral representing the total attenuation of the corresponding X-rays as they cross the scanned object.

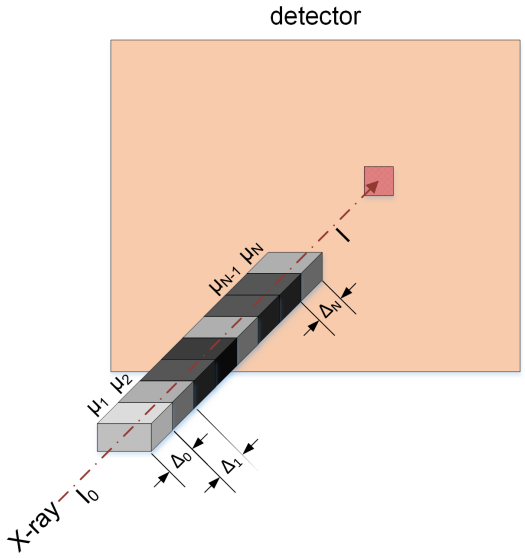


Figure 3.14.: A detector element (pixel) measures the remaining X-ray intensity after being attenuated by the materials (μ_i crossed by the said X-ray. The X-ray has an initial intensity I_0.

The attenuation of X-rays passing through the scanned object is defined by Beer's law as following:

$$I = I_0 exp(\sum_i (-\mu_i \Delta_i)) \tag{3.1}$$

where I_0 and I are the initial and final X-ray intensities, respectively. The term μ_i is the material linear attenuation coefficient, and Δ_i is the intersection length of the X-ray beam with each of the said materials having the coefficients μ_i. During the acquisition of the X-ray images, required for the CBCT reconstruction, the detector measures the values I. The value I_0 is usually measured during the detector calibration by acquiring X-ray images such that no object is scanned (only air). This is repeated for different exposures in order to determine an appropriate gain correction in order to keep the detector response in a linear range, as depicted in Figure 3.15. The maximum gray value corresponds to I_0.

The Equation 3.1 can be rewritten by applying a logarithmic scaling as

$$\ln(I_0/I) = \ln(G_{max}/G) = \sum_i \mu_i \Delta_i \tag{3.2}$$

where G_{max} is the maximum gray value and G the recorded pixel gray value. The equation 3.2 gives the linear relationship between the CT volume and the X-ray data: each detector measurement (pixel value) can be considered (after logarithmic re-scaling) as the sum of all attenuation coefficients μ_i along the projection line attaining the said pixel, i.e.

$$\ln(G_{max}) - \ln(G) = \mu_1 + \mu_2 + \cdots + \mu_N \tag{3.3}$$

For illustration, the dark pixels in the raw X-ray image shown in Figure 3.16(a) correspond to the X-rays which are highly attenuated by bone structures, i.e. the recorded intensity is low. In contrast, the same pixels appear as bright pixels in the logarithmic scaled X-ray images (b), since they represent the sum of the (high) attenuation coefficients of the bone structure along the corresponding projection lines.

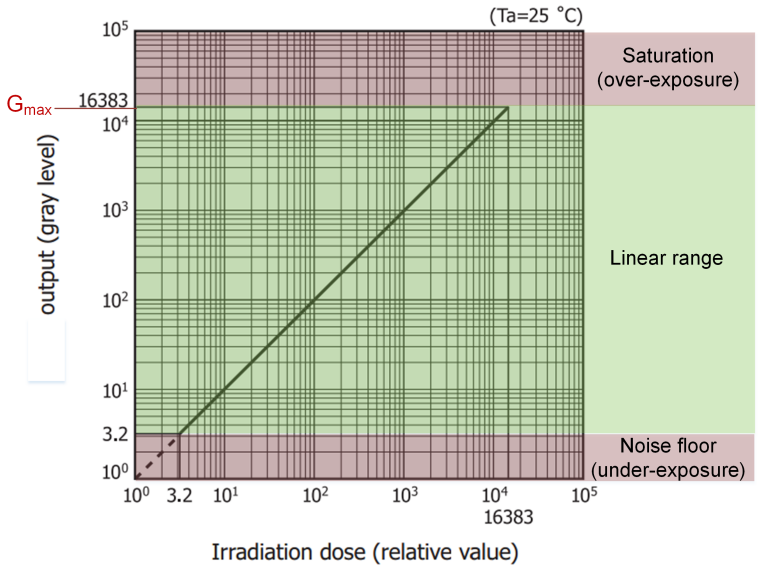
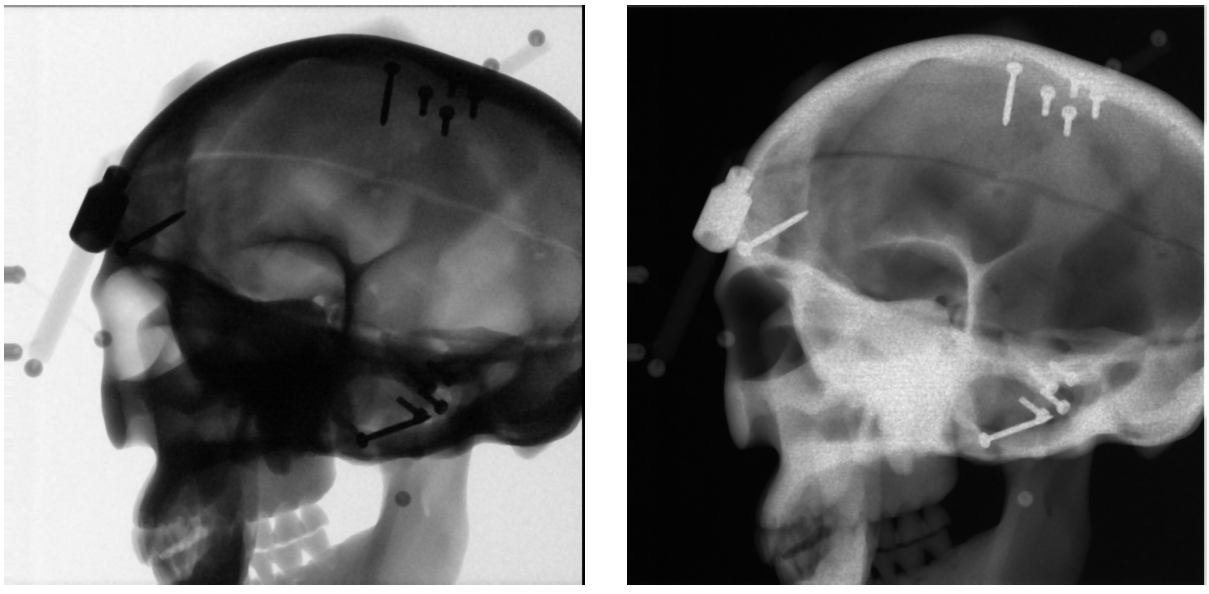


Figure 3.15.: Typical example of output linearity of 14-bit flat panel detector [ham].



(a) (b)

Figure 3.16.: An example showing an X-ray image before (a) and after (b) applying the logarithmic scaling. The later (b) is provided as output for the CBCT reconstruction.

In this chapter we refer to the images gained after the logarithmic scaling according to Equation 3.3 as X-ray images P_θ. These are provided as input for the reconstruction process.

The equation 3.2 can be formulated as

$$P_\theta(\vec{r}) = P_\theta(q) = \int V(\vec{s}_\theta + t\vec{r})dt \tag{3.4}$$

where $\vec{s}_\theta$ is the X-ray source position at the gantry angle θ, and $\vec{r}$ is the X-ray direction. Based on the

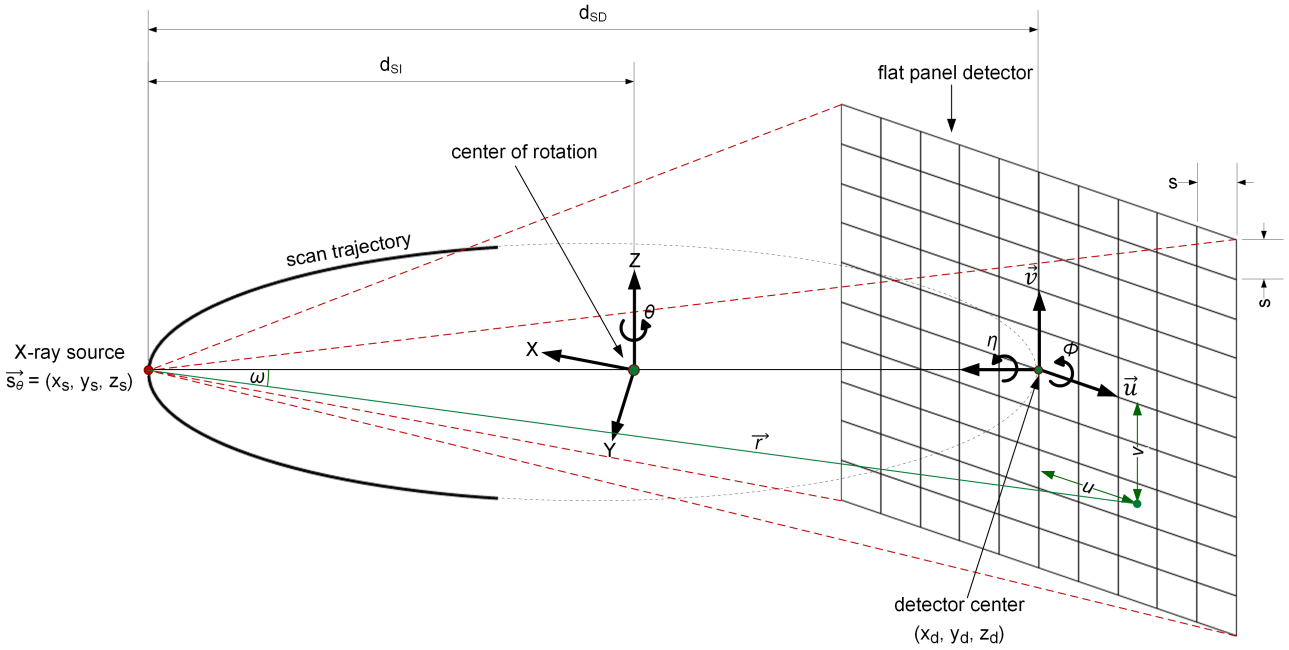


Figure 3.17.: Schematic diagram of the CBCT projection geometry.

equation 3.4, the reconstruction problem can be formulated as an inverse problem aiming to estimate the CT volume V from the observations P. In the following, conventional reconstruction algorithms serving to solve this inverse problem are presented and discussed. For a detailed description of the different reconstruction techniques the reader is referred to [KS01, Her09].

3.3.2. Analytic reconstruction techniques

3.3.2.1. Fourier-based reconstruction

By exploiting the analogy between the Radon transformation and the Fourier transformation, the central slice theorem can be formulated and used for 3D reconstruction [KS01, Tuy83]. However, this theorem holds only for parallel-beam projections. For fan-beam, as well as CBCT projections, a rebinning has to be applied to the projection data in order to sort it into an equivalent parallel-beam projection data. Usually this step corrupts the input projection data with additional errors, which may obviously degrade the reconstruction quality. Furthermore, the Central-slice theorem gives samples on a polar grid. However, standard inverse Fourier transform –applied to reconstruct the volume– requires data placed on a rectangular grid. Therefore the data has to be resampled to a Cartesian grid, which is associated with additional

interpolation error resulting in typical CT artifacts such as blurring and ringing. Some remedies have been presented to attenuate this problem as proposed in [O'S85, JMNM91].

3.3.2.2. Filtered Back-Projection

Feldkamp, Davis, and Kress adapted the Filtered Back-Projection (FBP) algorithm [BR67], developed to reconstruct 2D images, for processing cone-beam data also from partial circle trajectory [FDK84]. This algorithm has been successfully implemented on the most of medical CT imaging systems basically due to its computational efficiency. In the following, the different processing steps of the FDK algorithm are listed. A derivation of the algorithm is given in [FDK84].

1) Cosine weighting. In order to normalize the distance between the detector cell and the X-ray source position, each pixel value of the X-ray images P_θ is weighted as following:

$$P_\theta^{(1)}(\vec{r}) = P_\theta^{(1)}(q) = P_\theta^{(1)}(u,v) = \frac{d}{\sqrt{u^2+v^2+d^2}} P_\theta(u,v) \tag{3.5}$$

where $q = (u,v)^T$ are the coordinates of the projection of the voxel i onto the detector plane and θ the gantry angle. The weighting factor $\frac{d}{\sqrt{u^2+v^2+d^2}}$ is equal to the cosine of the angle ω, as depicted in Figure 3.17.

2) Data redundancy weighting. In case of partial circle trajectory, some line integrals may be measured twice. Thus, a data redundancy weighting is performed as

$$P_\theta^{(2)}(u,v) = w_\theta(u) \cdot P_\theta^{(1)}(u,v) \tag{3.6}$$

where the weighing function $w_\theta(u)$ (known as Parker weighting [Par82]) is defined for circle trajectory as

$$w_\theta(u) = \begin{cases} \sin^2\left(\frac{\pi\theta}{4(\omega_{max} - \arctan(\frac{u}{d}))}\right) & \text{if } 0 \le \theta \le 2\omega_{max} - 2\arctan(\frac{u}{d}) \\ 1 & \text{if } 2\omega_{max} - 2\arctan(\frac{u}{d}) \le \theta \le \pi - 2\arctan(\frac{u}{d}) \\ \sin^2\left(\frac{\pi(\pi+2\omega_{max}-\theta)}{4(\omega_{max}+\arctan(\frac{u}{d}))}\right) & \text{if } \pi - 2\arctan(\frac{u}{d}) \le \theta \le \pi + \omega_{max} \end{cases} \tag{3.7}$$

To be noted is that the weighting function is independent of the v-coordinates, because the points outside the trajectory plane are treated based on their orthogonal projection onto that plane.

3) Ramp filtering. Assuming that the plane containing the scan trajectory is parallel to $\vec{u}$, the Ramp filtering is applied following this equation

$$P_\theta^{(3)}(u,v) = \int_{-\infty}^{+\infty} \underbrace{\frac{-1}{2\pi^2(u-u')^2}}_{\text{Ramp kernel}} P_\theta^{(2)}(u',v)du' \tag{3.8}$$

An unwanted effect of ramp filter is the amplification of noise. In case of real measurements, this problem can be overcome by applying a mixture of a ramp-filter and low-pass filter like the Hanning filter [OSB99]. The main idea behind is to cut-off high frequencies characterizing the noise.

4) Back-projection. The voxel values of the object under investigation are recovered by back-projecting the filtered projections $P_\theta^{(3)}$ into the 3D image space according to this equation:

$$V(i) = d_{SD} \cdot d_{SI} \int_{\theta min}^{\theta max} \frac{P_\theta^{(3)}(q)}{||\vec{r}||} d\theta \tag{3.9}$$

where $\vec{s}_\theta = (x_s, y_s, z_s)$ is the X-ray source coordinates, d_{SD} the distance from the X-ray source to the detector, and d_{SI} the distance from the source to the iso-center.

Limitations. The FDK algorithm have several limitations [SHG*11, MDNK10]:

- it requires ideally circular scan trajectory. Usually, this requirement can not be satisfied in case of C-arm systems due to large deviations (several millimeters) from the ideal path.
- it expects linear response of the detector elements. This assumptions is obviously not fulfilled if metal parts are scanned. Even in ideal conditions a linear response can be approximated only in a small dynamic range.
- it can not handle large cone angles: an exact reconstruction can only be achieved in the center slice.

These main drawbacks make the FDK method not appropriate for the investigated CBCT system.

3.3.3. Iterative reconstruction techniques

Iterative reconstruction (IR) methods have several advantages compared to direct analytical methods. For instance, they can incorporate prior knowledge about the detector response, system geometry, scanned object, etc. The IR methods solve the inverse problem using algebraic approaches. They define therefor a set of linear equations. Then the reconstructed image is iteratively determined by solving this system of equations. As depicted in Figure 3.18, Each iteration involves two main steps: the forward- and back-projection of an intermediate reconstruction.

The forward-projection step simulates synthesized projections of the intermediate reconstruction called Digitally Reconstructed Radiographs (DRRs). By comparing the DRRs with the acquired X-ray images, correction terms are computed. These terms are used to update the intermediate volume in order to reduce the discrepancy between the simulated projections and the acquired data. The iterative reconstruction is finished either when a fixed number of iterations is reached, or the correction term is small, i.e. the process converges towards a stable solution.

In the following a review of the most popular IR algorithms are presented. For more details, the reader is referred to [BKK12].

Several iterative reconstruction algorithms have been proposed. These methods can be classified in two main categories:

- algebraic reconstruction algorithms:
  - algebraic reconstruction technique (ART) [GBH70, KS01, Col77, FST*13]
  - simultaneous algebraic reconstruction technique (SART) [KS01, KP09]
  - multiplicative algebraic reconstruction technique (MART) [KS01, KP09]
- statistical reconstruction algorithms:
  - statistical iterative reconstruction technique (SIRT) [Col77, KS01, WLL14]

- maximum likelihood expectation-maximization (MLEM) [MGK*95]
- iterative coordinate descent (ICD) [BDLT10]

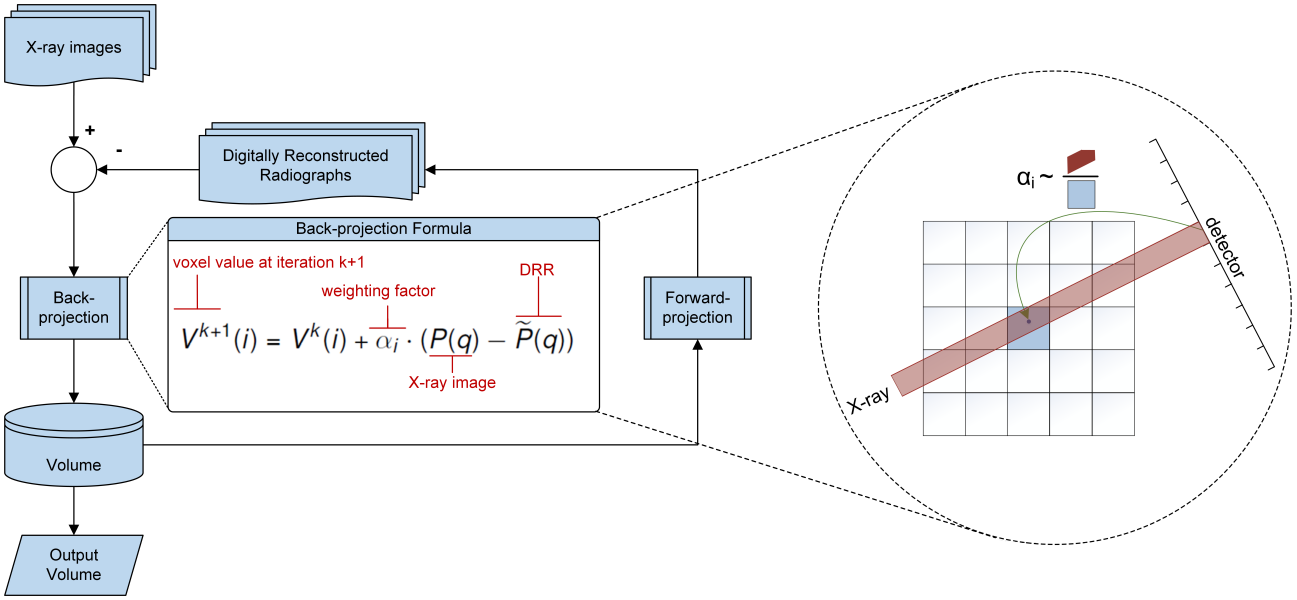


Figure 3.18.: The typical pipeline of iterative reconstruction algorithms.

Problem formulation. Given the following linear system

$$\mathbf{A}V = P \tag{3.10}$$

where $\mathbf{A} = \left[a_{iq}\right]$ is the matrix mapping the CT voxels $V(i)$ to the projection pixels $P(q)$ with a_{iq} the contribution of the voxel at position i to the pixel at position q. A CT volume V is lexicographically ordered as a column vector of size $(N \times 1)$, with N the number of voxels to be reconstructed. Similarly, P is a $Q \times 1$ column vector composed of all projection values for all pixels over all gantry angles, i.e. the pixels of all X-ray images. The matrix $\mathbf{A}$ is therefore of size $(Q \times N)$. Each image voxel at position i corresponds to one column of the system matrix A. Similarly, each projection ray, i.e. a single pixel q, corresponds to one row of $\mathbf{A}$. For instance,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{11} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{Q1} & a_{Q2} & \dots & a_{QN} \end{bmatrix}, \qquad V = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_N \end{bmatrix}, \qquad P = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_Q \end{bmatrix}.$$

Direct methods, such as Cholesky Factorization, are not practical to solve this problem due to the potential lack of sparsity of the coefficient matrix $\mathbf{A}$ and due to its large size. Alternatively, algebraic and statistical methods have been proposed to solve the mentioned object function in an iterative way.

3.3.3.1. Algebraic reconstruction algorithms

One of the first applied iterative reconstruction methods is the algebraic reconstruction technique (ART) [GBH70]. It aims at solving the mentioned linear system (Equation 3.10) by applying the Kaczmarz method [Kac37]. All iterative reconstruction techniques consist of two main operations to update the current CT volume estimate, namely the forward- and the back-projection (see Figure 3.18). The forward projection crates DRRs $\tilde{P}$, whereas the back-projection updates voxels values by back-projecting a correction term along a given projection line. These two operators are describe with more details below.

The algorithm starts with an initial guess $V^{(0)}$, e.g. an initial reconstruction with FDK algorithm. At each iteration step k, one measured projection value (pixel) $P(q)$ is selected from the X-ray images and compared to the corresponding projection value $\widetilde{P}(q)$ computed based on the last volume estimate $V^{(k-1)}$ by forward-projecting $V^{(k-1)}$ on the position q (see below). Then the difference $(P(q) - \widetilde{P}(q))$ is calculated and distributed equally among all voxels along the corresponding ray $\vec{r}_q$: The value of each voxel $V(i)$ contributing to the pixel $P(q)$ is updated by adding a correction term which is proportional to the difference $(P(q) - \widetilde{P}(q))$ and the factor a_{iq} (see below). By iterating this process, the current estimate $V^{(k)}$ is expected to get closer to the unknown CBCT volume to be recovered.

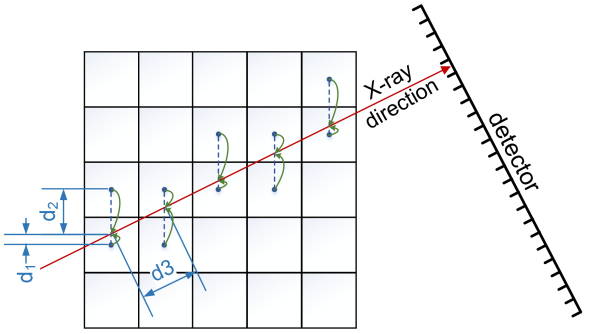


Figure 3.19.: Illustration of the forward projection: a line integral is computed for each point detector (pixel) by summing all voxel attenuation values along this line (X-ray direction).

Forward projection. Forward projection of a known CBCT volume simulates a virtual X-ray image called Digitally Reconstructed Radiograph (DRR) by calculating the line integrals of the said volume for each point on the detector. Several methods have been proposed to achieve this task. They can be classified in voxel-driven and ray-driven techniques [BS04] . A simple and computationally efficient method is the algorithm proposed by Joseph [Jos82]: in order to compute the line integral through a CT along a particular line (X-ray) $\vec{r}_q$, the sum of all voxel attenuation values along this line is calculated. This is done by performing these steps:

1. select all voxels traversed by a ray $\vec{r}_q$
2. perform a linear interpolation between adjacent voxels as shown in 3.19, i.e. based on the distances d_1 and d_2
3. multiply the interpolated value (from step 1) with the intersection length d_3
4. perform the steps 1 and 2 for all intersection lengths along $\vec{r}_q$, and sum the calculated values

Compared to other techniques, e.g. the method proposed by Siddon [Sid85], the Joseph's method is computationally efficient and does not suffer from aliasing artifacts. However, is still inaccurate, since it interpolates linearly between only 4 adjacent voxels in 3D space, the interpolated value is kept constant over the full distance d_3. Better results can be achieved by:

- performing a tri-linear interpolation in order to get a better approximation of the voxel value at the position of interest
- reducing the step length d_3. From our experience, reducing it to the half can enhance the DRR quality without causing significant additional computational cost
- calculating the exact the intersection volume within each voxel as proposed in [Sak93].

Back-projection. The reverse (adjoint) operation of the forward projection is the back-projection, so that voxels traversed by a given X-ray are updated with weighted contributions a_{iq} from the pixel corresponding to the said ray. Typically, the contribution is weighted based on the (normalized) intersection length of the projection line with the voxel of interest. An illustration of this operation is shown in Figure 3.20.

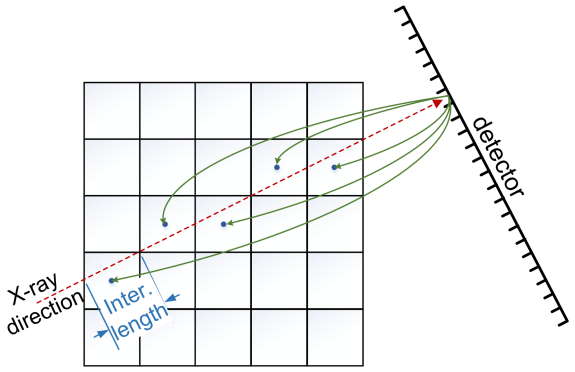
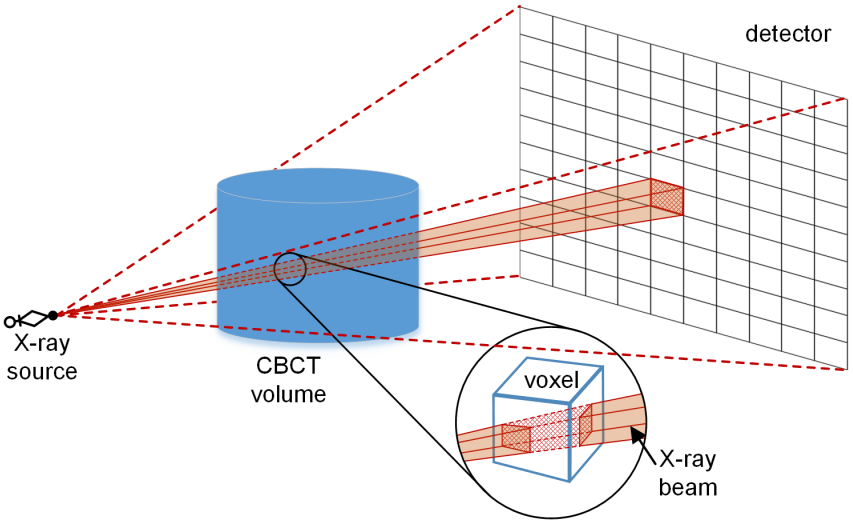


Figure 3.20.: Illustration of the back-projection operator: values of all voxels traversed by a particular X-ray corresponding to a certain detector pixel are updated. The correction terms are typically set proportional to the (normalized) intersection lengths of the voxels with the current X-ray. Green arrows show the contribution factors a_{iq}.

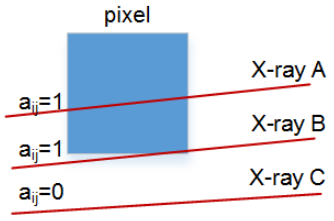
Obviously, the weight coefficients a_{iq} are the elements which links the CT voxel values to the X-ray values (pixels). For a good reconstruction quality, it is mandatory that each element a_{iq} do accurately represent the contribution (influence) of the ray $\vec{r}_q$ passing through the pixel q on the voxel at position i. Ideally, the contribution (weight coefficient) should be proportional to intersection volume between the cone beam (pyramid of X-rays) captured by a given detector element and a voxel traversed by this beam, as illustrated in Figure 3.21(a). However, applying such a model is computationally expensive. In the first ART method proposed by Gordon et al. [GBH70], the weights a_{iq} was set to 1 if a ray $\vec{r}_q$ touch or pass through the voxel i, otherwise it was set to 0 (see Figure 3.21(b) for 2D case). A more accurate approach was proposed by Shepp and Logan [SL74]: a_{iq} was set proportional to the area of intersection of the voxel i with $\vec{r}_q$, simulated as a ray beam not a line, as depicted in Figure 3.21(c). An efficient and accurate alternative is to set the weights a_{iq} proportional to the intersection length of the ray $\vec{r}_q$ (simulated as line) with the voxel i [Sid85], as illustrated in Figure 3.21(d).

Mathematically, an approximation of the solution of the linear system formulated in Equation 3.10 can be expressed for each voxel $V(i)$ at position i as

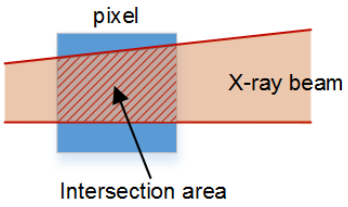
$$V^{(k+1)}(i) = V^{(k)}(i) + \lambda \frac{P(q) - \sum_{j=1}^{N} a_{jq} V^{(k)}(i)}{\sum_{j=1}^{N} a_{jq}^2} a_{iq} \tag{3.11}$$



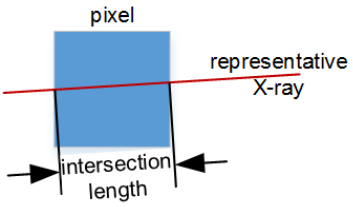
(a) Intersection volume of an X-ray beam (bundle of X-rays) with a voxel.



(b) binary contribution: 1 in case of intersection, otherwise 0



(c) contribution set proportional to the intersection area



(d) contribution set proportional to the intersection length

Figure 3.21.: Illustration of the intersection of an X-ray beam with a volume voxel in 3D (a) as well as in 2D (b,c,d). The contribution should be set proportional to the intersection volume between an X-ray beam and a given voxel (a). Three models which approximate this are shown in (b,c,d).

where λ is a relaxation parameter (see next page), N is the number of all voxels, k the iteration number and a_q^T the transposed row of the system matrix $\mathbf{A}$ corresponding to the pixel at position q.

Unlike ART which operates on single pixels, i.e. single rays, the simultaneous algebraic reconstruction technique (SART) updates the complete X-ray image P_θ, i.e. all pixels corresponding to the gantry angle θ. This can be mathematically expressed as following

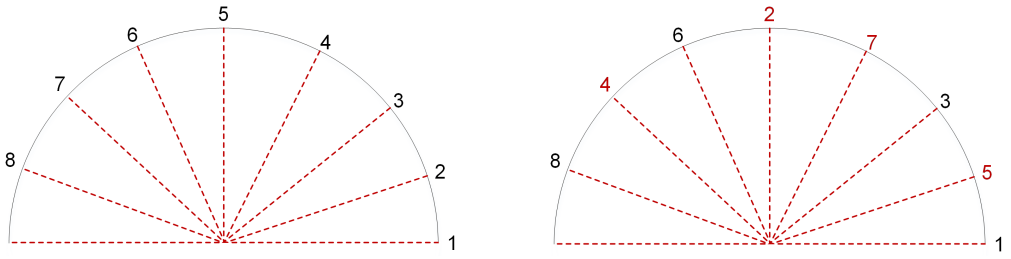
$$V^{(k+1)}(i) = V^{(k)}(i) + \lambda \frac{\sum\limits_{q \in P_\theta} \left[\frac{P(q) - \sum\limits_{j=1}^{N} a_{jq} V^{(k)}(j)}{\sum\limits_{j=1}^{N} a_{jq}^2} \right] a_{iq}}{\sum\limits_{P(q) \in P_\theta} a_{jq}} \tag{3.12}$$

Gorddon et al. [GBH70] have also proposed a multiplicative algebraic reconstruction technique (MART) which performs the update by multiplying the current voxel estimate with the correction term, unlike ART

which adds the correction term to the current solution. The update equation is expressed as

$$V^{(k+1)}(i) = \lambda \frac{P(q)}{\sum_{j=1}^{N} a_{jq} V^{(k)}(j)} V^{(k)}(i) \tag{3.13}$$

Additive ART may update the voxels values to be negative and have therefore to be corrected at the end stage of the process e.g. by stetting the negative values to zero. In contrast, MART has the advantage of reconstructing only positive values. Moreover, additive ART tends to minimize the overall variance of the CT reconstruction while MART maximizes the entropy [Cri04]. According to [KS01], additive techniques are more susceptible to additive noise than the MART, however the influence of this aspect on the reconstruction quality is very limited. Usually, both techniques have similar reconstruction quality such that the difference can not be perceived by visual inspection.

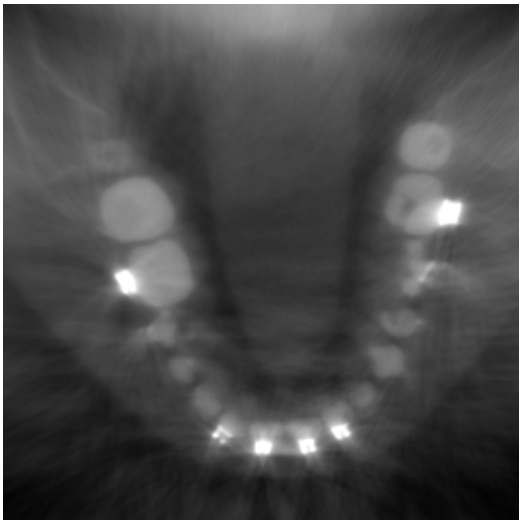


(a) acquisition order (b) re-ordered images

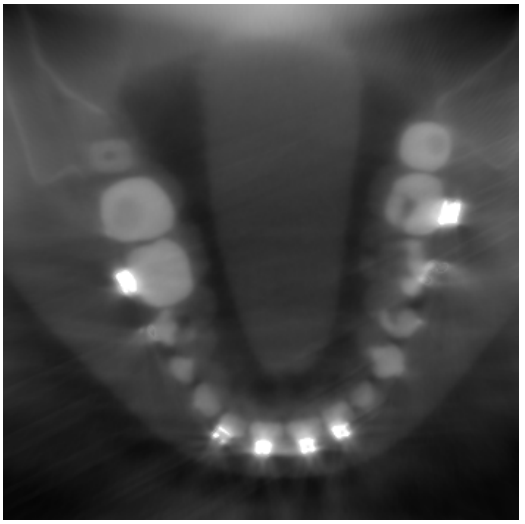
Figure 3.22.: Two possible access orders for the X-ray images. By changing the order, the correlation between consecutively processed X-ray images is minimized, which increases the convergence rate.

Factors which influence the reconstruction quality. Several factors have significant impact on the reconstruction quality when iterative reconstruction algorithms are applied:

- the accuracy of the weighting factors a_{iq}: obviously more accurate weighting factors would result in better reconstruction quality, since the real imaging process would be better approximated by the imaging model.
- the order in which the X-ray images are processed: appropriate order, e.g in form of ordered subsets (OS), would reduce the number of required iterations and improve the rate of convergence [HL94]. For instance, an example is shown in Figure 3.22. Following the acquisition order shown in Fig.3.22(a), consecutive projections are highly correlated. In contrast, if the order is adjusted as shown Fig.3.22(b), the correlation of the consecutive projections is expected to be minimal, which would speed up the convergence. This is illustrated in Figure 3.23. We refer to [HR94] for more details.
- the initial guess $V^{(0)}$: it can influence the final solution to which the method converges.
- the relaxation factor λ (also called step length): it controls the convergence of the method. If this value is high ($\gg$1), the method converges fast but may result in divergent high frequency artifacts. In case of very small λ (close to zero) the method converges slowly, however the reconstructed images would be very smooth. For good reconstruction results and convergence rate, the relaxation factor should be set in the range of [0.5, 1.0].



(a) the X-ray images are processed in the same order as in the acquisition



(b) re-ordered X-ray images in order to minimize the correlation between consecutively processed images

Figure 3.23.: An example illustrating the influence of the order in which the X-ray images are processed on the convergence rate: the CBCT volume was reconstructed twice with the same number of projections and iterations; only the order in which the X-ray images were processed was changed. The volume reconstructed with re-ordered X-ray images (b) has better quality compared to the volume reconstructed while processing the X-ray images in the same order in which these images were acquired.

In order to illustrate the problem of conventional back-projection kernels while dealing with detector metal artifacts we refer to the example shown in Figure 3.24. For simplicity we assume that each voxel is projected on only one pixel, and that all voxels along a given projection ray are updated by the same correction term computed for the said ray. In this example, we initialize a CT volume by setting its voxel densities to the correct values. In this case, if the reconstruction method is accurate and robust, it should not worsen the current volume estimate. We distinguish between two cases:

- Linear response: The detector operates in its linear range so that the recorded value at each pixel is equal to the sum of attenuation coefficients of all voxels crossed by the corresponding projection ray. In this ideal case the created DRRs are identical to the X-ray images, since they are created based on the correct CT volume. Consequently the computed correction terms (difference between a DRR and the corresponding X-ray image) are equal to zero, thus the current (correct) voxel values remain the same after each iteration.
- Non-linear response: As illustrated by the intensity profiles (X-ray images P_A and P_B) acquired at two different gantry angles (see Figure 3.24), the values recorded by the detector pixels on which metal implants are projected are inferior to the ideal (correct) projection value (illustrated by the green lines). Thus, wrong correction terms are typically computed at these (metal) pixels, which would result in metal artifacts, as described below.

Let focus on three representative voxels (see Figure 3.24):

- a metal voxel $V(M)$ which has an initial value equal to 4000,
- a soft tissue voxel $V(S)$ whose initial value is 200,
- and a soft tissue voxel $V(A)$ which would be reconstructed as artifact. Its initial value is 200.

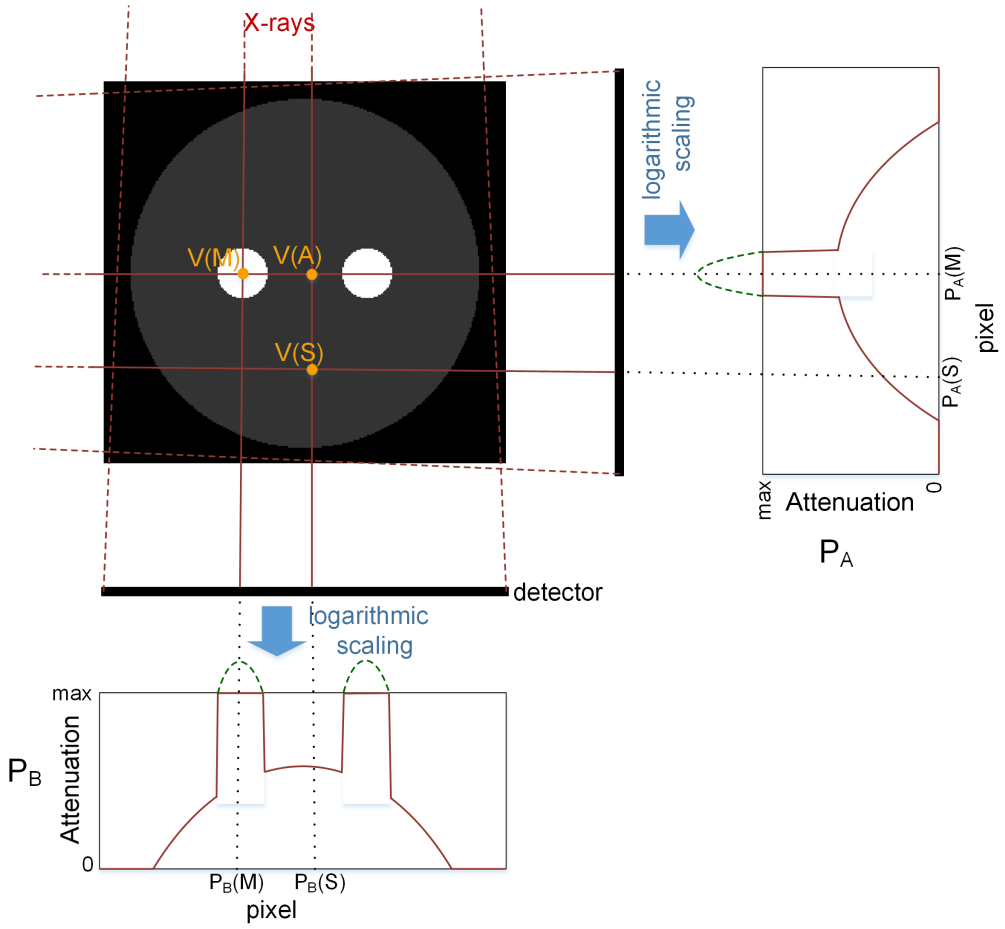


Figure 3.24.: An example illustrating why metal artifacts are created. The X-rays images of a double-spheres phantom are not ideal: the detector response at pixels where metal implants are projected is not proportional to the sum of attenuation coefficients of voxels crossed by the corresponding projection rays.

The update of the values of the considered voxels is done as follows:

1. Starting by the X-ray image P_A:
   - The artifact-free soft tissue voxel $V(S)$ is projected on the pixel $P_A(S)$. The difference between the X-ray image and the corresponding DRR created based on the current (correct) CT volume is computed at the said pixel. The calculated difference is equal to zero, since all voxels have the correct value, and the detector response is in the linear range. Consequently, the correction term to be back-projected is zero, thus the voxel $V(S)$ keeps its current (correct) value.
   - The metal voxel $V(M)$ is projected on the pixel $P_A(M)$ where the detector response is not linear (see noise floor in Figure 3.15). Thus, the recorded value is inferior to the sum of all attenuation coefficients along the projection ray captured by the said pixel. The difference between the DRR and the X-ray image computed at this pixel position is therefore negative, which results in a negative correction term (calculated by dividing this difference by the total number of voxels crossed by the projection ray). This correction term is small compared to the current value of the metal voxel, so that the change of $V(M)$ is also relatively small. It is clear that the voxel value is falsified after performing this step, however this is not significant, since the relative change is negligible (around 1%).

- The artifact voxel $V(A)$ is projected on the same position as the metal voxel $V(M)$, i.e the pixel $P_A(M)$. Thus, it is updated by the same correction term which is applied to $V(M)$, however, this contribution is significant compared to the current voxel value $V(M)$ which would be reduced by around 20%. Similarly, the densities of all voxel lying between the metallic spheres are also reduced by around 20% from their current values.

2. Processing the X-ray image P_B:
   - By forward projecting the new volume estimate, the value of the DRR pixel $\widetilde{P}_B(S)$ would be inferior to the expected value $P_B(S)$, since several voxels which are projected on the said pixel were updated in the previous step to have values smaller than their correct values. Thus, a positive difference would be computed for this pixel. By dividing this difference by the number of all voxels along the corresponding projection ray, a relatively small positive correction term is calculated. When this correction term is added to $V(S)$, its current value is falsified, which is also the case all other voxels crossed by the said projection ray.
   - The voxel $V(A)$ is also projected on $P_B(S)$, thus it is also update by the same correction term similar to $V(S)$. However, this positive correction is too small to compensate the inappropriate negative contribution from the previous back-projection of the X-ray image P_A.
   - Similar as by processing the X-ray image P_A, the metal voxel $V(M)$ is again updated by subtracting a relatively small contribution compared to its current value.

It is clear that all voxels are deviated from their correct values after only one iteration. Performing further iterations based on the current falsified CT volume would only amplify the current error. It is to be noted that such errors are more perceptible at soft tissue voxels, since the relative change there is more significant in comparison to metal voxels.

The aspect mentioned above is the same when the CT voxel values are set to zero. For instance, a representative CT slice is shown at several stages of the reconstruction process in Figure 3.25. The mentioned problem get more an more excessive with the increasing number of processed X-ray images.

3.3.3.2. Statistical reconstruction algorithms

The detected photons are expected to follow a Poisson distribution. This assumption is not satisfied for very low dose due to the influence of electronic noise [YLL*09], which is especially the case in positron emission tomography (PET) and single-photon emission computed tomography (SPECT). One of the most successful statistical reconstruction method is the expectation-maximization (EM) proposed by Shepp and Vardi [SV82] for emission tomography.

Some statistical methods try to consider the counting statistics of the detected photons by processing the raw data projections (X-ray images) [La 05] or the reconstructed CT images by applying denoising techniques in order to reduce the noise while preserving spatial resolution. These algorithms can therefore be classified as pre- and post-processing approaches. Due to the performed denoising, high-frequency details might be lost depending on the strength of the filtration method.

More sophisticated statistical techniques try to reduce the noise by modeling the statistics of photons – during the acquisition– and incorporate this model in the reconstruction process. They perform usually based on maximum likelihood (ML) [MGK*95] or the least squares (LS) principles, such as the iterative coordinate descent (ICD) algorithm [BDLT10]. However, only an approximated model can be employed

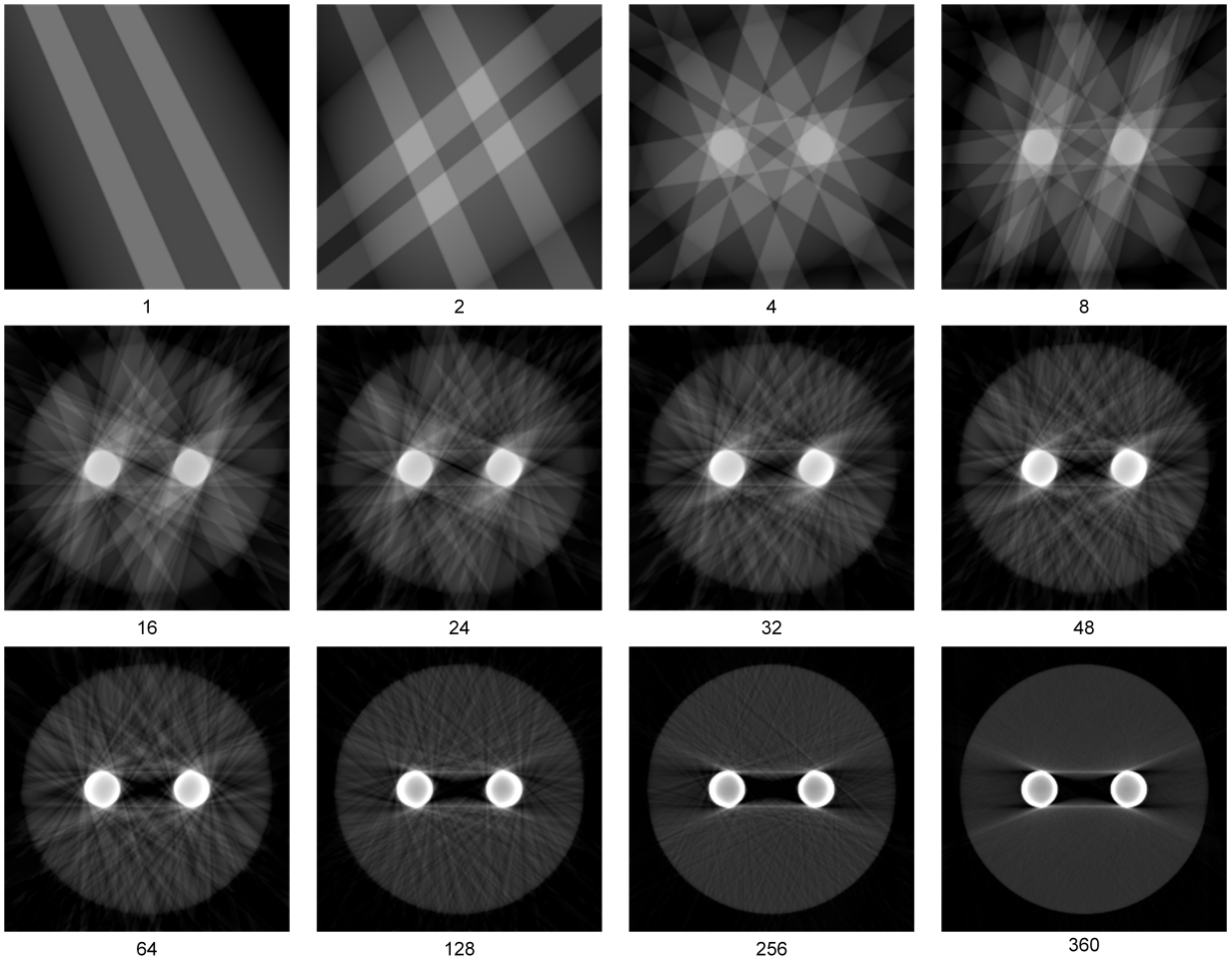


Figure 3.25.: An example showing the evolution of metal artifacts in a representative CT slice during the reconstruction of a CT volume by applying the SART algorithm.

due to the complexity of the physical X-ray projection phenomena. In addition, statistical methods impose a regularization term which aims at suppressing the noise and CT artifacts while preserving the high-frequency details. Many regularization models have been proposed to impose e.g. local smoothness on the reconstructed image [EF99]. Therefore, a tradeoff between the data fidelity and the regularization term, i.e. noise-resolution tradeoff, has to be found.

In order to enhance the convergence speed, the ordered subsets (OS) has been proposed [HL94]. It should lead to faster convergence by dividing the projection images into sub-groups called subsets. The update is then performed for each subset instead of the complete projection set. By combining known algorithms with OS, new methods have been proposed, such as OS-SIRT, OS-ICD, and OS-EM [MGK*95].

Compared to ART methods, statistical methods employ more accurate physical models compared to ART, which would result in better reconstruction quality. However, they require much more processing time and have usually several tuning parameters which should be adjusted based on the scanned object, exposure level, etc. For this reasons, statistical methods are only conditionally recommended for clinical application.

3.3.4. Comparison of analytical and iterative reconstruction techniques

Iterative reconstruction techniques outperform the analytical approaches in term of reconstruction quality, especially in case of noisy or incomplete projections. Effectively, analytical methods ignore the measurement noise in the problem formulation and solve the noise-related problems with post-filtering operations, which usually results in reconstruction quality degradation. In addition, iterative reconstruction techniques can operate at reduced patient dose because they can precisely model the acquisition process [BH07].

A certain disadvantage of IR methods is their high computational complexity. Unlike analytical algorithms which are performed in a single iteration, iterative reconstruction techniques are executed in an iterative scheme, which results in higher computational demand. Above this, one iteration is more expensive as two FDK reconstructions. Only through high computational capacities offered by modern computers, such as graphics processing units (GPUs), it is possible to reconstruct high-resolution CTs ($> 512^3$ voxels) with IR methods in an acceptable time range (< 1 minute).

A comparison between analytical and iterative reconstruction techniques is summarized in table 3.1. The used terms are briefly explained here:

- Simplicity: indicates the theoretical complexity of the reconstruction algorithm
- Prior knowledge: if it is possible to incorporate prior-knowledge, e.g. about the scatter, in the reconstruction process
- Computational cost: refers mainly to the running time
- Reconstruction quality: correctness of the reconstructed volumes, i.e. less CT artifacts are created.

	Analytic Reconstruction Techniques	Iterative Reconstruction Techniques
Simplicity	✓✓	✓
Prior knowledge	✗	✓✓
Computational cost	✓✓	✗
Reconstruction quality	✓	✓✓

Table 3.1.: Comparison between analytical and iterative reconstruction techniques.

3.3.5. Metal artifact reduction techniques

Metal artifacts, appearing in form of streaks and shadows in the CT image as shown in section 3.2.4, are caused by multiple mechanisms related to high-density objects (with high atomic number), like metal implants. The presence of such objects cause beam hardening and scatter effects, which typically results in dark streaks –between highly attenuating structures– with surrounding bright streaks.

Several metal artifact reduction (MAR) techniques have been presented in the last decade. They can be classified in two categories: CT- and X-ray-based methods.

The CT-based methods are directly applied to the reconstructed CT volume [ZRW*00, KIF*07, HBB*06]. Kennedy et al. [KIF*07] proposed to segment the defective voxels and re-estimate their values using pattern recognition methods. A knowledge-based MAR method has been introduced by Hammil et al [HBB*06]. Their method modifies the corrupted volume through a sequence of morphological operations and 3D filtering. All of these techniques result usually in an artificial appearance of the artefactual regions

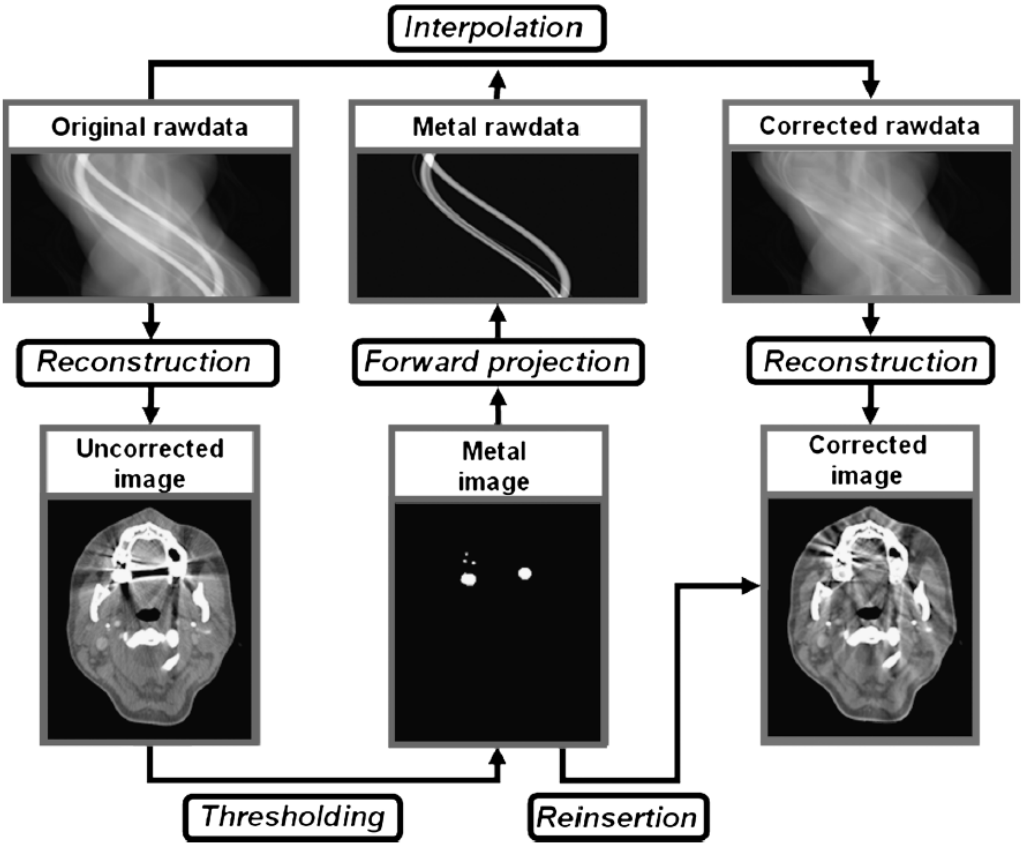


Figure 3.26.: Process chart of the metal reduction algorithm proposed by Lell et al. [LMK*12].

and the surrounding tissues in the reconstructed volumes respectively slices. In addition, structures in corrupted regions can not be recovered.

Better results can be achieved by X-ray-based methods [KHE87, MRL*10, VJvG10, LMK*12, WLZ*13]. They consider the projection data generated by beams through high-density objects as missing measurements. These approaches consist of two main processing steps prior to the final image reconstruction: segmentation of metallic parts and completion of X-ray images. The segmentation of metal objects is performed in the image domain, i.e. based on the CT slices of the initial reconstruction, and its result is then used to determine the regions in the X-ray images which have to be modified. This step aims to replace the metal pixels in order to generate a metal-free X-ray images, e.g with simple interpolation or inpainting techniques (see below).

Usually the segmentation is performed by simple thresholding [LMK*12]. Despite its simplicity, relatively accurate results can be achieved since metal objects have extremely high CT values compared to other materials. However, in case of complex structures, like dental braces, streaking artifacts which have intensities similar to metal voxels are also erroneously identified as "'metal voxels"'. As claimed in several studies, any segmentation error may induce additional artifacts in the whole reconstructed image [MRL*10, LMK*12]. More sophisticated segmentation techniques have been proposed but most of them assume a Hounsfield Unit (HU) calibration [MRL*10, BS06], which is not fulfilled in most of CBCT C-Arm systems.

The replacement of metals pixels can be performed with basic interpolation techniques, like linear interpolation, can be applied. This could be sufficient for small metal objects, but not for big objects since new secondary streak artifacts would be generated. In order to overcome this limitation, completion of X-ray images employing Priors or Image Inpainting have been proposed. Bal and Spies [BS06] suggested to estimate missing data using the forward projected information from predefined tissue-class model. Meyer et al. [MRL*10] used sinograms of prior images to enhance the homogeneity of the recovered regions. Thereby inaccurate priors can lead to quality degradation in the recovered sinogram and hence to substantial secondary artifacts in the corrected image [KCWM12]. Furthermore, optimal parameter tuning is required to achieve good results. It has been shown that the inpainting techniques outperform the other methods, but it is associated with a significant increase in computational cost [ZPH*11].

The performance and robustness of the conventional MAR techniques is still questioned, especially when several metal implants with unknown geometry are scanned. The conventional MAR methods which are based on segmentation of metallic parts cannot avoid segmentation errors because streak artifacts typically appear as bright as metal and are therefore mostly wrongly segmented. The rest of MAR reconstruction techniques cannot significantly reduce the metal artifacts or result in synthetic appearance of the reconstructed CT volume. A more robust method which can completely eliminate the metal artifact is therefore required.

3.4. Concept of CBCT improvements

In this chapter, novel approaches aiming to deal with some common CBCT reconstruction artifacts are presented. Some of the limitations addressed in Section 3.2.4 can be relatively easily solved by enhancing the acquisition device, e.g in case of:

- limited gantry angle: extend the scan trajectory such that the gantry angle get over 180° plus fan angle
- scatter: employ anti-scatter grids to attenuate this problem
- noise: improve the detector response as well as the calibration process, such as the calibration of uncalibrated or defective detector cells

Other limitations can be solved by the reconstruction algorithm. For instance, it is not possible to acquire the theoretically required minimum number of X-ray images, since the number of projections should be kept as low as possible in order to minimize the radiation exposure. However, it is well known that iterative reconstruction techniques can achieve good reconstruction quality even in case of moderate number of X-ray images (≈150) [KS01].

Other limitations are more challenging and require more sophisticated techniques to deal with:

1. **Mechanical inaccuracies.** The mechanical inaccuracies can not be entirely solved with pre-calibration, since this technique can only compensate systematical (repetitive) errors. In order to compensate stochastic errors (vibrations) occurring during the scan process, online calibration should be applied. Such a calibration requires usually additional hardware and excessive processing time, which make it impractical for clinical application. Therefore, a new on-the-fly (online) calibration is proposed to compensate minimal stochastic inaccuracies occurring during the acquisition of the X-ray images. The novel method, described in details in Section 3.5, consists of the following steps:

a) Perform an initial estimate of the CT volume using the original X-ray images and the input (pre-calibrated) geometrical parameters
b) Generate digitally reconstructed radiographs (DRRs) according to the correct projection parameters
c) Align the originally acquired X-ray images to the simulated DRRs
d) Update the geometric projection parameters of each X-ray employing the alignment matrices from the third step
e) Reconstruct a new CT volume based on the improved geometric parameters
f) Perform the steps 2 to 5 iteratively.

The proposed method makes use of an intensity-based 2D-3D registration, namely in the third step (c), which is susceptible to CT artifacts. In order to overcome this problem a robust similarity measure has been developed. It accounts for CT artifacts in order to reduce their influence on the registration result.

2. **Iterative VMR-based CBCT reconstruction**. Iterative CBCT reconstruction methods consist basically of two main steps: forward and back-projections. Unlike the first step which is relatively straightforward, the back-projection is more complicated and has a major impact on the reconstruction quality. Several CT artifacts are caused by the reconstruction algorithm itself due to the met simplifications, e.g in modeling the physical acquisition process. As it was shown in Section 3.3.3, conventional reconstruction techniques cannot deal with the non-linear detector response, which would result in excessive CT artifacts. In this case, better distribution of the correction terms along the corresponding projection rays should enhance the reconstruction quality. Accordingly, a new non-linear back-projection kernel is proposed in Section 3.6 to improve the CBCT quality by exploiting information present in the variance-to-mean ratio map (presented in Section 3.5.3). The proposed kernel makes use of the distribution of all projected values corresponding to a given volume voxel in order to reduce the background artifacts level and enhance the soft tissue contrast. An explicit segmentation of the artifact regions or a modification of the X-ray images are therefor not required.

3. **Metal artifact reduction (MAR).** Conventional techniques dealing with this issue have extra tuning parameters which have to be adequately set, else new reconstruction artifacts would be generated. Robust algorithms, which do not require extra tuning or prior-knowledge about the scanned object, have to be developed in order to ensure good reconstruction quality and high practicability in clinical applications.

   A new iterative reconstruction method which aims to enhance the reconstruction quality and reduce (and even eliminate) the metal-related artifacts is proposed in Section 3.7. It consists of the following steps:
   a) Reconstruct an initial CBCT volume based on the acquired X-ray images
   b) Detect the “metal voxels” (high-density voxels) in the CBCT volume and compute their corresponding projection areas within the X-rays
   c) Generate metal-free X-ray images by replacing “metal pixels” in the acquired X-ray data
   d) Reconstruct a second CBCT volume based on the estimated metal-free X-ray images computed in the previous step
   e) Reconstruct a third (final) CT volume by fusing the volumes reconstructed in the first and fourth steps

The second step requires accurate segmentation of metal objects, since segmentation errors would cause new excessive CT artifacts. Therefore, a novel fully automatic segmentation method is proposed to achieve this task. It does not require prior knowledge about the metallic objects present in the scanned region.

These three main topics (limitations) are the center focus of this section. Some contributions of this chapter have been presented in SPIE Medical Imaging 2014 [ES14] and will appear in [ES15].

Assumptions and simplifications

An accurate modeling of the physical acquisition system should substantially enhance the reconstruction quality and minimize the CT artifacts. However, it is not practicable to try to account for all physical effects due to the associated high computational complexity. The goal of this section is to present simplifications which do not result in substantial degradation in quality while keeping the computational complexity in acceptable level. These simplifications are defined according to actual standards, with focus on the C-arm CBCT system under investigation. It is to be noted that while performing the back-projection the volume update is following a voxel-driven scheme, this means that we iterate over all volume voxels. For each voxel, the projection of its center point is computed and then the pixel value corresponding to the calculated projection coordinates is back-projected. Similarly, in order to generate a DRR, we iterate over all pixels.

The focal spot of the X-ray tube (X-ray generator) as well as the detector elements (pixels) are modeled as points. In reality the focal spot and the pixels are areas with given extent. The used C-arm system, as well as similar current systems, have a small focal spot size of 0.3 mm as well as small pixel size of 0.283×0.283 mm^2 with a source-detector-distance of 1110 mm. Therefore, it is justifiable to model them as points with zero extent.

The number of the measured photons at the detector –after the attenuation of the X-rays by the scanned object– are assumed to follow a Poisson distribution. This assumption is violated in case of very low dose due to the influence of electronic noise. The investigated C-arm CBCT system operates in the range of 40 to 120 kV. In such a case the measured photons can be assumed to be Poisson distributed [GBD04, SHG*11]. A better approximation is the log-likelihood for compound Poisson statistics [WME*06], but it is impractical for reconstruction.

In the system under investigation, a pre-filtration of generated X-rays with a physical device (filter made of 0.1 mm Cu) serves to ensure that the x-ray beams approximated a truly mono-energetic beams. Consequently, it is reasonable to assume mono-energetic X-ray beams in case of the investigated C-arm system since it is equipped with such a physical filter.

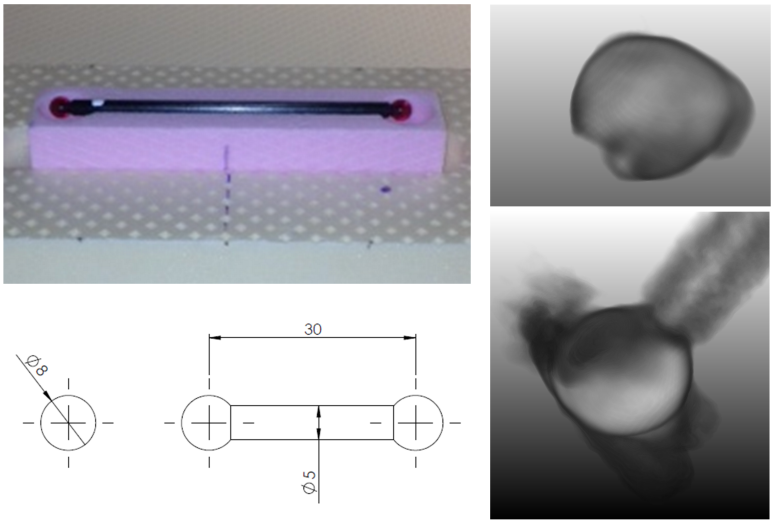
In case of excessive scattering effect where photons leave the expected straight ray, the reconstruction quality may be substantially degraded if this aspect is not considered. The used C-arm CBCT system is equipped with an anti-scatter grid (pb 8/70) which prevents the detection of scattered photons with large deflection angle. It was shown that this measure is able to decrease the image noise by up to 70% [SSZ*12]. An explicit modeling of scatter, e.g. with Monte Carlo methods, is therefore not mandatory, especially that such simulations require high computation times even on modern GPUs.

The inhomogeneity problem which can be caused by the automatic dose reduction is not relevant in case of head CT, since the total X-ray attenuation is more or less the same over all gantry angles, so that the required exposure is kept more or less constant. This means that the selected gain is kept constant during the acquisition of the X-ray images. In order to be sure that this is really the case, the automatic dose control unit was deactivated, i.e. the exposure was enforced to be constant over all projections.

3.5. On-the-fly geometric calibration

Mobile C-arm CBCT systems have an open design and are therefore not as stable as spiral CT systems. Effectively, the motion of the X-ray source and the detector may differ significantly from the ideal scan trajectory. A precise estimation of the geometrical projection parameters is essential to perform an accurate 3D reconstruction, otherwise reconstruction artifacts and loss of details are expected. In order to illustrate the effect of mechanical inaccuracies on the reconstruction quality, a double-ball-bar phantom is scanned by our C-arm CBCT system. As shown in Figure 3.27, the reconstructed balls do not have a spherical shape due to the mechanical inaccuracies since the projection images do not exactly match the estimated geometrical projection parameters.

Fortunately, the C-arm motion has a long-term reproducibility (over a period of 6 months) [DSC*08], so that a frequent pre-calibration is not necessary. Therefore, an offline calibration can be performed in order to compensate the mechanical inaccuracies by computing the deviation of the gantry and the detector from their intended ideal positions. However, this still be insufficient to achieve high-accurate reconstructions since it cannot compensate stochastic misalignments occurring during the acquisition of the X-ray images.



(a) Double-ball-bar phantom (b) Reconstructed balls

Figure 3.27.: Example illustrating the influence of geometric inaccuracies on the reconstruction Accuracy. Due to geometrical misalignments, the reconstructed balls loss their spherical shape.

Several online and offline calibration approaches have been suggested in the literature, but all of them require a phantom of known geometry [DSC*08, SH12] or an external tracking system [CBW*07, MNS00]. A phantom consists usually of precisely placed high-density spheres embedded in low attenuation substrate material, so that the spheres can be easily localized in the X-ray images (see Figure 3.28). After determin-

ing the coordinates of the spheres in the acquired projections, and given that their correct 3D positions in the original phantom are known, it is possible to estimate all geometrical parameters. A detailed review of calibration methods is presented in Section 4.2.

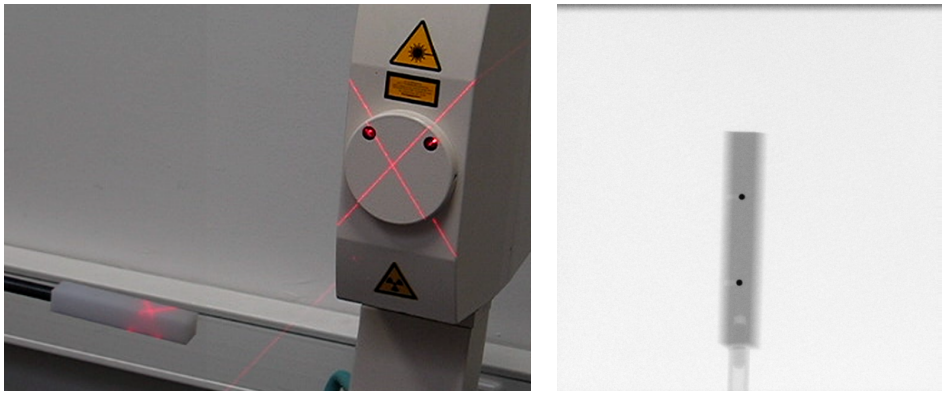


Figure 3.28.: A calibration phantom (left) and a corresponding 2D X-ray image (right) are shown. The metallic spheres appears as dark points with high contrast in the acquired scan and can therefore be easily localized and employed for geometrical calibration.

3.5.1. Concept

This thesis contributes a new on-the-fly (online) geometric calibration algorithm which aims to compensate stochastic-mechanical misalignments without additional phantoms or hardware. The proposed method uses the CBCT reconstruction itself, varies its corresponding projection parameters and improves the reconstruction in an iterative way. Therefore, neither phantom nor specific scan procedure is needed. Due to its computational efficiency, the proposed algorithm can be applied on-the-fly intra-operatively for removing CT reconstruction artifacts caused by mechanical inaccuracies. Two main assumptions are met to successfully perform the proposed calibration method:

- The remaining error after performing a pre-calibration (offline calibration) is minor ($\leq$ 1mm), which is a standard in current C-arm systems. The proposed algorithm can therefore be considered as a fine-tuning of the performed pre-calibration
- The object to be scanned is of rigid anatomy, e.g. head. This assumption serves to guarantee high registration accuracy. Indeed it is also possible to deal with deformable anatomies, but in this case the registration results may easily fail to achieve the desired sub-pixel accuracy due to the high number of unknown parameters to be estimated. This assumption is satisfied in the investigated case of lateral skull base surgery.

The underlying idea of the proposed approach is to perform an initial estimate of the CT volume and then generate digitally reconstructed radiographs (DRRs) according to the projection parameters for the original X-ray acquisitions. The originally acquired X-ray images are then registered to the DRRs, i.e. to the initial CT volume. In the ideal case, registration of DRR to X-ray images must be perfect, i.e the registration-transformation matrix **T** between the DRRs and the acquired X-ray images is equal to the identity matrix, otherwise the geometric projection parameters are updated based on the calculated transformation (**T**). A new volume is reconstructed using the improved geometric parameters and the original X-ray images, then the process is performed again. The pipeline of this iterative approach is shown in Figure 3.29 and illustrated in Figure 3.30.

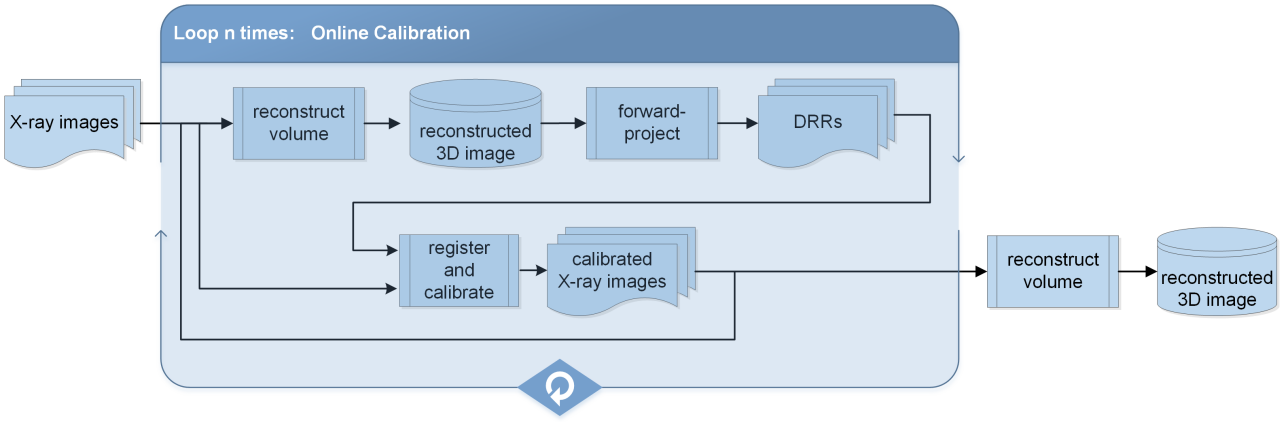


Figure 3.29.: The pipeline of the proposed on-the-fly (online) geometrical calibration method.

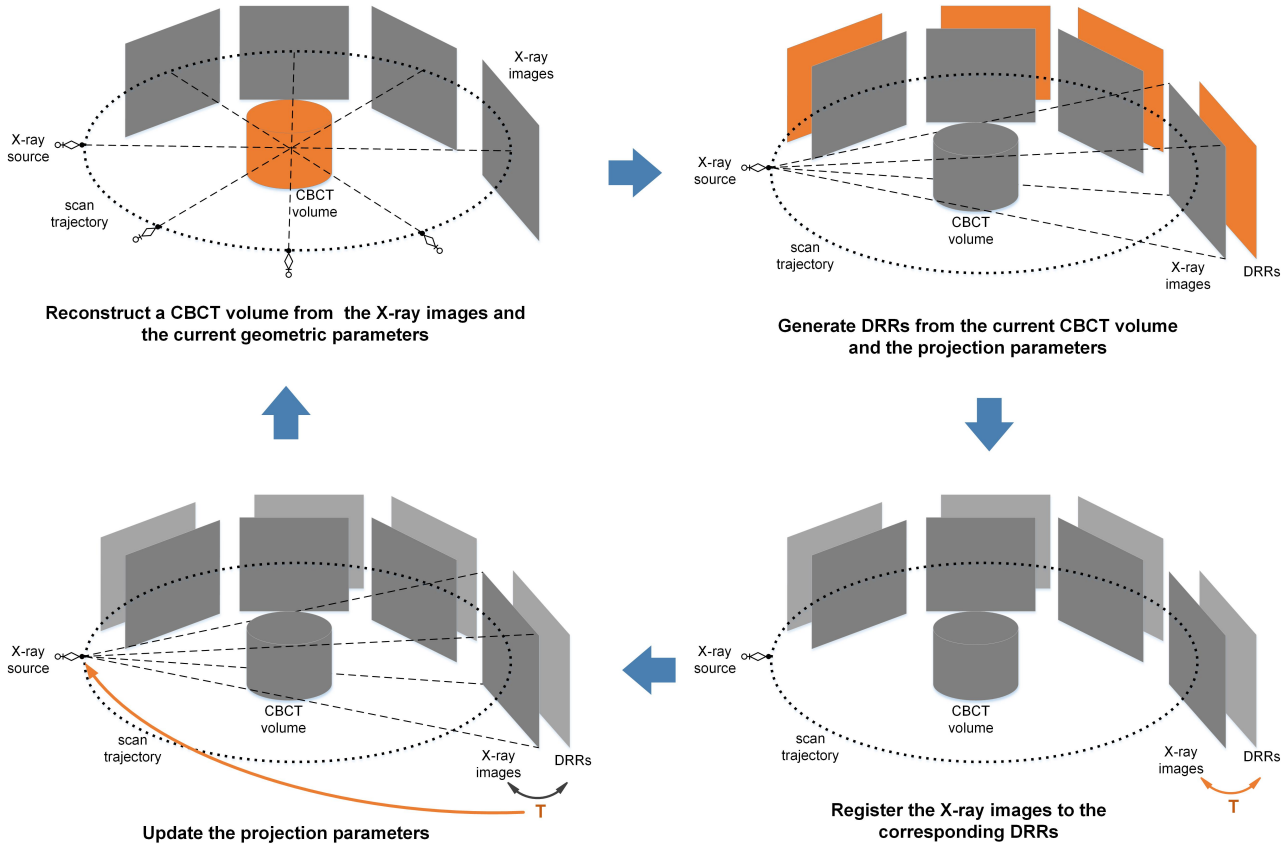


Figure 3.30.: An illustration of the proposed on-the-fly (online) geometrical calibration method.

The system geometry of the CBCT C-arm is shown in Figure 3.17. Given that the imaging system is modeled as a pinhole camera, the calibration process aims at determining a 3×4 homogenous projection matrix $\mathbf{H}_\theta$ which maps 3D volume voxel coordinates (x, y, z) to the corresponding 2D image pixel coordinates (u, v) on the X-ray image P_θ:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{H}_\theta \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{3.14}$$

Using all matrices $\mathbf{H}_\theta^{(0)}$ and the X-ray images P_θ, an initial CT volume $V^{(0)}$ is reconstructed. DRRs are then simulated based on $V^{(0)}$ and $\mathbf{H}_\theta^{(0)}$. After performing a pre-calibration, each DRR $\widetilde{P}_\theta$ or X-ray image P_θ is associated with an initial projection matrix $\mathbf{H}_\theta^{(0)}$. In every iteration k of the online calibration, the C-arm pose (position and orientation relative to the patient) is re-estimated for each X-ray image. In order to achieve this task, the current CT volume is registered to each acquired X-ray image P_θ by employing an intensity-based 2D-3D rigid registration (see below).The output of this registration $\mathbf{T}_\theta^{(k-1)}$ is employed to update $\mathbf{H}_\theta^{(k-1)}$ in order to get to be a better estimate $\mathbf{H}_\theta^{(k)}$ of the projection matrix for the X-ray image P_θ and the k-th iteration:

$$\mathbf{H}_\theta^{(k)} = (\mathbf{T}_\theta^{(k-1)} \cdot \mathbf{T}_\theta^{(k-2)} \cdot \mathbf{T}_\theta^{(k-3)} \ldots \mathbf{T}_\theta^{(1)} \cdot \mathbf{T}_\theta^{(0)}) \cdot \mathbf{H}_\theta^{(0)} \tag{3.15}$$

Using P_θ and $\mathbf{H}_\theta^{(k)}$, a new CT is reconstructed, which is a better approximation of the original volume, and the process iterates again. In our experience 3 to 5 iterations are sufficient for a good result.

C-arm pose estimation. Given a CT volume and an X-ray image, it is to determine the correct position and orientation of the X-ray image (i.e. the detector) relative to the volume. If the computed pose is correct, a DRR created at this pose would match perfectly with the X-ray image.

In case of rigid anatomy, such as in this thesis, the transformation matching each CT voxel with the corresponding X-ray pixel can be described with a (perspective) projection matrix $\mathbf{H}_\theta$:

$$\mathbf{H} = \mathbf{K} \cdot \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_1 \\ r_{2,1} & r_{2,2} & r_{2,3} & t_2 \\ r_{3,1} & r_{3,2} & r_{3,3} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.16}$$

The matrix $\mathbf{H}$ can be spitted into extrinsic ($\mathbf{R}_{ot}$ and $\mathbf{T}_{rans}$) and intrinsic ($\mathbf{K}$) matrices. The extrinsic matrix defines the position and orientation of the C-arm in form of a 3×3 rotation matrix and a 3×1 translation matrix. The 3×3 intrinsic matrix $\mathbf{K}$ comprehends the detector pixel size s , the source-detector distance d_{sd}, the skew γ, and the projection of the isocenter on the detector plane (u_o, v_o):

$$\mathbf{K} = \begin{bmatrix} d_{SD}/s & \gamma & u_o \\ 0 & d_{SD}/s & v_o \\ 0 & 0 & 1 \end{bmatrix} \tag{3.17}$$

Since the intrinsic parameters are usually known and remain constant during the acquisition, the calibration task can be reduced to estimate the extrinsic parameters, i.e. the C-arm pose relative to the scanned object (patient).

The pose estimation can be achieved with a 2D-3D registration which is performed as follows (see Figure 3.31):

1. create a DRR for each of the X-ray image based on the current CT volume
2. compute the similarity score (metric) between each DRR and the corresponding X-ray image. This similarity measure tells how well the X-ray and the DRR match to each other
3. vary the pose parameters in order to maximize the similarity score, i.e to calculate a DRR which matches perfectly with the current X-ray image. Therefore, the similarity score is provided as input for an optimization algorithm which aims at altering the projection parameters to maximize the similarity score (expressed as a cost function)
4. iterate over steps 1-3 until finding the best pose.

The registration is terminated when changes in the projection parameters became less than 0.01mm for translational parameters and 0.01° for rotational parameters.

In this thesis a variety of similarity measures (metrics) as well as optimization algorithms has been proposed. The similarity measures are described with more details in Section 3.5.3. For an overview about the different optimization algorithms, the reader is referred to [vDH*11]. It has been shown in several publications that the Powell-Brent method [Pow64] is a reliable optimization method for the problem under investigation in term of accuracy and computational complexity [Ste14, vDH*11, dPAB90]. Furthermore, this method has the largest capture range compared to other optimization techniques. Therefore, the Powell-Brent optimizer was employed in this thesis to perform the intensity-based registration.

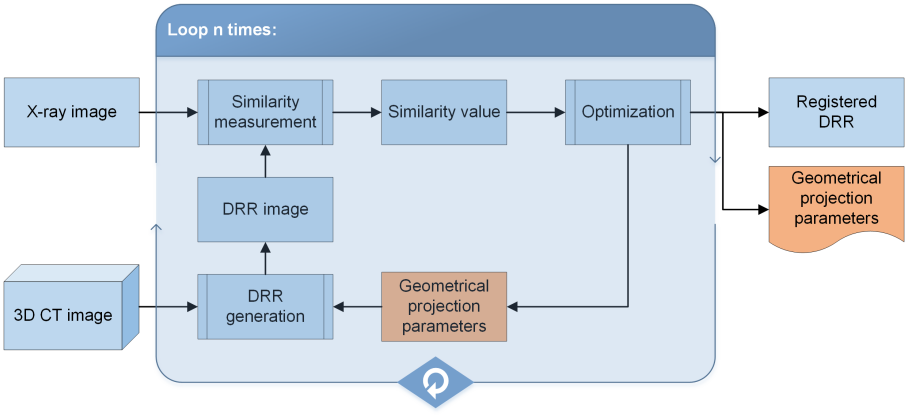


Figure 3.31.: Iterative 2D-3D intensity-based registration. The approach aims to determine the C-arm pose (geometrical projection parameters) by maximizing the similarity between the acquired X-ray image and created DRRs.

It is to be noted that the output of the offline calibration provides a good initial estimation of the C-arm pose. Thus, the registration process is initialized with the geometric parameters $\mathbf{H}_\theta^0$ provided by the pre-calibration. Similarly and for every iteration k, the registration is initialized with the projection parameters calculated in the last iteration $\mathbf{H}_\theta^{k-1}$ as seen in Equation 3.15.

3.5.2. Geometric parameters to be estimated

The nine geometrical extrinsic parameters being under instigation are: the source coordinates (x_s, y_s, z_s), the flat-panel detector center coordinates (x_d, y_d, z_d), and the detector orientation angles (ϕ, θ, η) relative to isocenter (see Figure 3.17).

Parameter	Symbol	Standard deviation	maximal deviation
X-ray source position	x_s	0.047 mm	0.19 mm
	y_s	0.018 mm	0.11 mm
	z_s	0.045 mm	0.20 mm
Detector position	x_d	0.047 mm	0.21 mm
	y_d	0.018 mm	0.12 mm
	z_d	0.046 mm	0.22 mm
Detector orientation	θ	0.007°	0.023°
	ϕ	0.003°	0.011°
	η	0.001°	0.0046°

Table 3.2.: Glossary of geometric parameters and their deviation after performing the pre-calibration.

Table 3.2 summarizes the measured geometrical deviations of the nine geometrical parameters of the studied scanner after performing the pre-calibration. This table was derived from the calibration protocol generated at the end of the offline calibration.

Performing the fine-tuning process while considering all nine geometrical projection parameters is time consuming and is therefore not practical for clinical applications. Thus, it is desirable to reduce the number of parameters to be estimated during the fine-tuning. A sensitivity study is therefore performed to select the geometric projection parameters having a large impact on the reconstruction accuracy. The imaging device used in this study is the C-arm CBCT system described in Section 3.2.1.

Sensitivity study. The influence of each of the nine projection parameters was separately evaluated as follows:

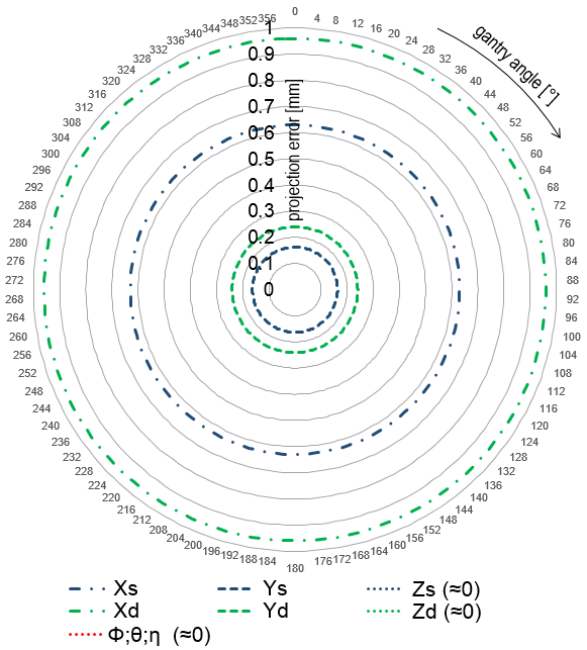
1. For each X-ray image P_θ select a projection parameter, e.g. x_s, the correct projection matrix $\mathbf{H}_\theta$ and neglect the contribution of the rest of parameters by setting them equal to the correct (original) values
2. Vary the value of the parameter under investigation by deviating it from its correct value by the maximum expected (once adding and once subtracting) error as resulting from the pre-calibration
3. Compute the resulting projection error after perturbing the parameter under investigation for every voxel and over all gantry angles.

For instance, in a simulated study let $\mathbf{H}_\theta$ be the ideal and correct projection matrix corresponding to the X-ray image P_θ, and let the parameter under investigation be x_s. The previous steps results in:

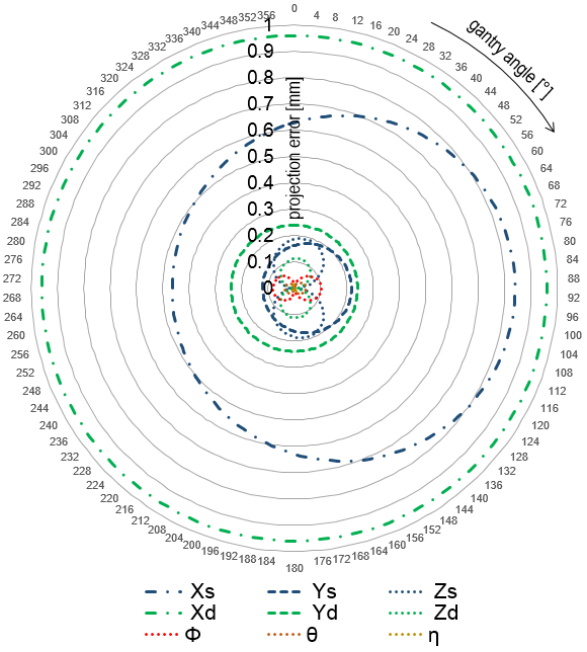
1. Setting the eight other parameters equal to their correct values
2. Corrupting the parameter x_s by its maximal expected error ($\pm$0.19 mm) and calculating the corresponding erroneous projection matrix $\widehat{H}_{\theta,x_s}$
3. Calculating the projection error ϵ_{x_s} for every voxel. ϵ_{x_s} is defined as the Euclidean distance $d(,)$ between its correct and its erroneous projection coordinates:

$$\epsilon_{x_s} = d(H_\theta \cdot \vec{x}, \widehat{H}_{\theta,x_s} \cdot \vec{x}) \tag{3.18}$$

   where $\vec{x} = (x, y, z, 1)$ are the voxel coordinates.
4. For each X-ray image P_θ, i.e. over all gantry angles, compute the maximum projection error for each projection parameter over all voxels and then display them over the gantry angles (see Figure 3.32).



(a) Voxel coordinates (0,0,0)



(b) Voxel coordinates (75,0,0)

Figure 3.32.: Projection errors induced by the maximum geometric non-ideality computed for each geometric projection parameter, namely the translational parameters (X_s, Y_s, Z_s, X_d, Y_d, Z_d) and the rotational parameters (ϕ, θ, η). The magnitude of the projected error depends mainly on the deviation from the reference geometric projection parameter, the gantry angle, and the location of the voxel to be projected. The projection error is plotted for two representative voxels.

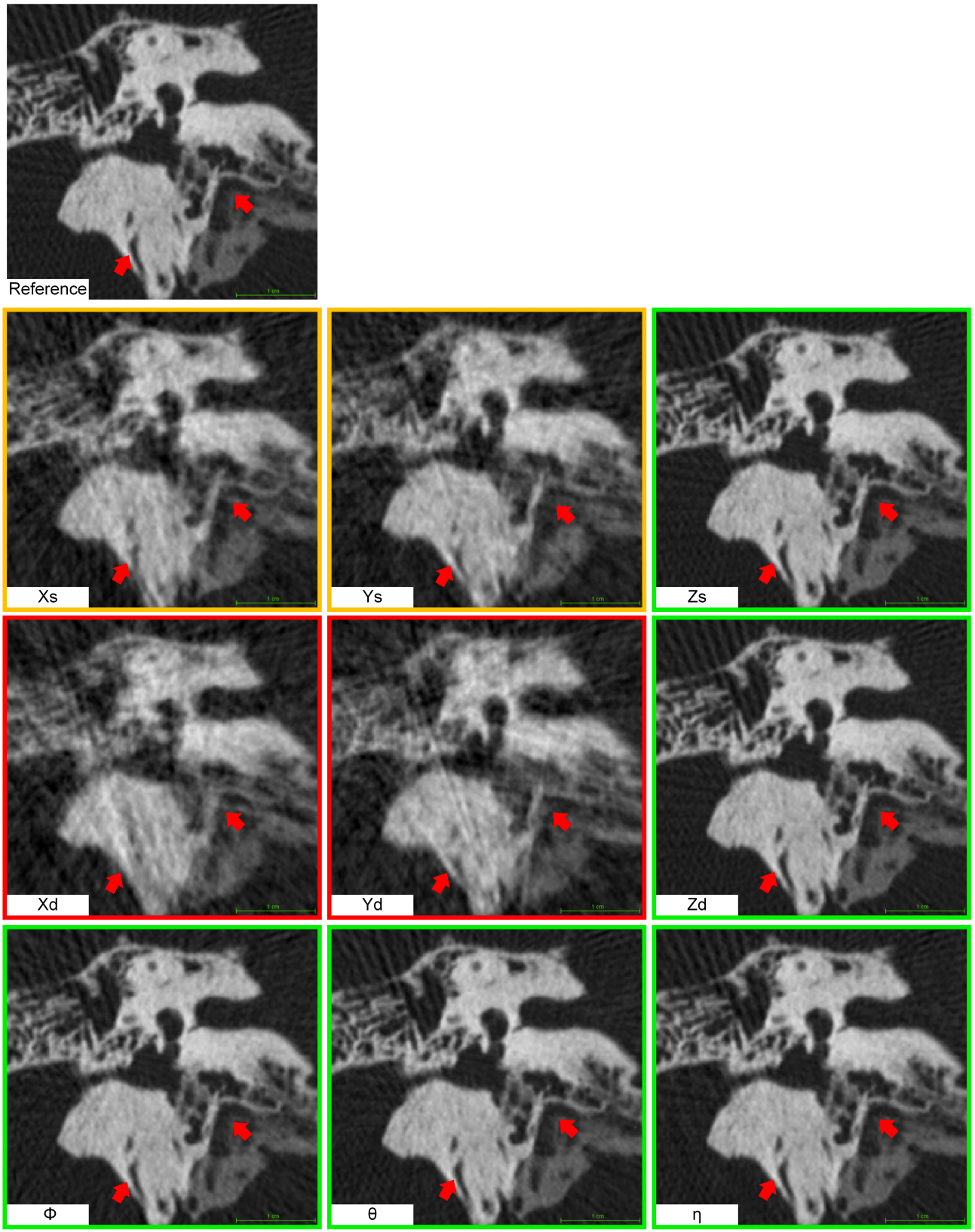


Figure 3.33.: Influence of inaccurate estimation of the geometrical projection parameters on the reconstruction quality. Translational parameters $(x_s, y_s, z_s, x_d, y_d, z_d)$ and rotational parameters (ϕ, θ, η) were deviated from their ideal values with Gaussian distributed error (μ = 0; σ_{trans} = 0.3; σ_{rot} = 0.03). Green, yellow, and red boxes refer to minor, moderate, and heavy degradation in the reconstruction quality, respectively.

As depicted in Figure 3.32, not all parameters show the same error amplification. In addition, the maximum projection error depends also on the location of the voxel to be projected. From these figures x_s, y_s, x_d, and y_d seem to have the highest influence on the error.

This was verified by visual inspection: starting with simulated or real CBCT acquisitions, we disturbed each geometric projection parameter individually and one-on-a-time by the maximum expected error from Table 3.2, produced a new CT volume and visually inspected the degradation in quality of the disturbed CT compared to the undisturbed reconstructed images. For this purpose, X-ray images with ideal (correct) projection matrices are required, which is easily achieved by simulated data as described in section 3.5.4, but cannot be provided in case of real data. Thus, we decided to scan a cadaver temporal bone with a spiral CT scanner (SOMATOM Definition AS+) in high-resolution mode and use the CT volume as "ground truth". The acquired CT volume is more realistic than mathematical models, since it reflects detailed anatomical features. Furthermore, spiral CTs show superior geometric stability and have geometric parameters which are substantially more close to the ideal ones compared to our CBCT scanner.

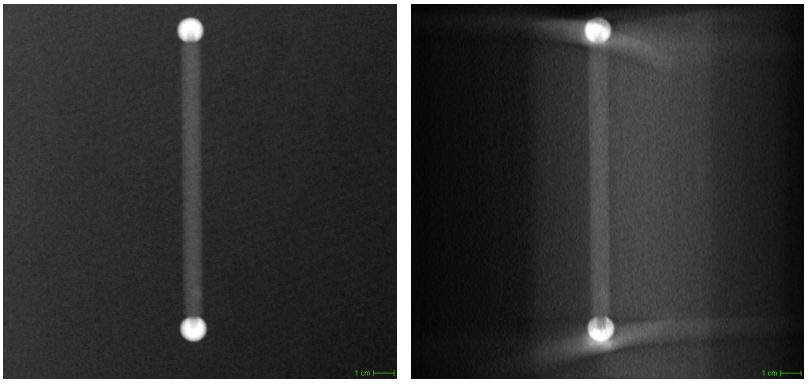
Based on the level of visual quality degradation, the parameters (x_s, y_s, x_d, y_d) are confirmed as main error sources since they have high visual influence on the reconstruction quality (see Figure 3.33). A deviation in x_s or y_s results in stretching along x or y directions, respectively. A similar effect is more dominant in case of erroneous y_d and y_d. The other parameters have also some influence on the reconstruction, but with limited degradation of the image quality, and can therefore be neglected.

In conclusion, it is not mandatory to account for all deviations from an ideal scan orbit. At least for the C-arm being under investigation, it is possible to reduce the 3D fine-tuning to 2D process considering only (x_s, y_s, x_d, y_d) without a substantial loss in accuracy or reconstruction quality.

This simplification is employed to reduce the computational time towards intra-operative, on-the-fly application, since only a single DRR and a rather trivial 2D-2D registration are required to correct the projection parameters of the corresponding X-ray image: In every k-th iteration of the online calibration, a 3×3 transformation matrix $\mathbf{T}_\theta^k$, which aligns the DRRs produced based on the CBCT of the k-th iteration to the acquired X-ray images, is computed by registering each P_θ to its corresponding DRR. The projection matrices are then updated based on the Equation 3.15.

3.5.3. Image registration

The task here is to register X-ray projections to its corresponding DRRs. Unlike the originally acquired X-ray images, which are artifact-free, DRRs calculated from reconstructed CBCT volumes may be corrupted by artifacts. This is the case if the CBCT volume, from which the DRRs are calculated, contains heavy degradations. Therefore, a main challenge within the proposed algorithm is to successfully align the DRRs to the acquired X-ray images, especially in the presence of CT inaccuracies such as metal-related artifacts. For instance, the DRR shown in Figure 3.34(b) exhibits dominant streak artifacts which might worsen the registration accuracy. In order to make the registration process robust against such artifacts, a weighted similarity measure (metric) which underweight the influence of the artifact regions on the registration result is proposed. Before describing the novel metric, the variance-to-mean ratio, which is employed in the proposed calibration process as well as in the CBCT reconstruction, is presented.



(a) Acquired X-ray image (b) DRR

Figure 3.34.: A representative X-ray image acquired with the investigated C-arm CBCT system (a) and a corresponding digitally reconstructed radiograph (b).

Variance-to-mean ratio. The variance-to-mean ratio (VMR) is a factor associated to each individual voxel. It is calculated as the variance of the intensities of all pixels lying on the projection paths of a certain volume voxel divided by the value of this voxel.

Given a voxel at position i with value $V(i)$, the center of this voxel is projected on each X-ray images P_θ at position $q_{i,\theta}$. The $VMR(i)$ is defined as

$$VMR(i) = \frac{\frac{1}{Q-1} \cdot \sum_{q=1}^{Q} \left(P_\theta(q_{i,\theta}) - \overline{p_i} \right)^2}{V(i)} \tag{3.19}$$

where $\overline{p_i}$ is the mean intensity of the pixels $P_\theta(q_{i,\theta})$ and Q is the total number of X-ray images.

The variance-to-mean ratio, also known as index of dispersion, dispersion index, coefficient of dispersion, or Fano factor [Hoe43], was used to quantify whether a set of measured/observed occurrences are clustered or dispersed compared to a standard statistical model. This thesis proposes to employ this ratio for CBCT reconstruction as a measure of the probability that a given voxel is affected by CT artifacts.

The VMR contains information about the spreading of the X-ray pixel values corresponding to a given volume voxel among all Q projections. If the VMR is equal to zero, the data is not dispersed. Such a case does not occur in real applications, even if the reconstructed volume is homogeneous, due to noise and differences in optical lengths (intersection lengths of the X-rays with the scanned object). If the value is smaller or equal to 1, the data is under-dispersed, e.g. in case of binomial or Poisson distribution. With increasing VMR, the projection values are more dispersed meaning that the confidence on these values should be questioned.

We can distinguish three cases:

1. Voxels with high attenuation (high-density voxels) tend to project on dark pixels on all Q projections, and have therefore low VMR values since all X-rays traversing these voxels are heavy attenuated, i.e. the variance of the pixel values among all images is small. By dividing this small variance by the high voxel density value, the resulting VMR gets even smaller.

2. Similarly, if a non-high-density voxel is not (or only occasionally) obscured by metallic objects, its corresponding pixels values on all projections are again similar to each other and their VMR will also be rather low.
3. The behavior described in case (2) changes significantly if metallic objects obscure a voxel in several (or many) projections: the corresponding pixels will be sometimes dark and other times gray, which results in high variance of their projected values among all Q projections. Divided by a low to moderate voxel density, the corresponding VMR is considerably higher than that of high-density or artifact-free voxels. In such a case the low-density voxel is usually erroneously reconstructed, typically as shadow or streak artifact. This is typically the case of voxels surrounding the metal voxels.

This thesis proposes to employ the VMR to enhance the calibration accuracy as well as the reconstruction quality.

1. **From the registration point of view:** the VMR can be considered as a measure to estimate the probability that a given volume voxel would belong to an artifact region. This probability is employed to enhance the registration accuracy required for the on-the-fly calibration
2. **From a reconstruction point of view:** the introduced VMR value serves to estimate the uncertainty of the reconstructed voxels caused by the spreading of the appropriate back-projection values. This is explained in the following: In case of a single circle scan trajectory, which is common in current C-arm systems, exact reconstruction algorithms as proposed by Defrise and Clack [DC94], become identical to FDK algorithm. In case of moderate cone-angles the FDK-reconstruction is a very good estimate of the scanned object if enough noise-free projections are provided and no metal objects are in the field-of-view. As shown in Section 3.3.2.2, a CT voxel value can be expressed as the mean of all corresponding filtered X-ray pixels (after applying a weighted filtering to the original projection data). Since the uncertainty in the mean value, i.e. the corresponding CT voxel value, is proportional to the standard deviation of the corresponding filtered X-ray values, it is reasonable to employ the variance as a measure for the the uncertainty of the estimated voxel value. A division by the mean value of the back-projection values, as it is the case in the VMR, serves for normalization.

However, applying the VMR as formulated in Equation 3.19 is associated with some drawbacks:

- the computed VMR-map is severely blurred since the high spacial frequencies lose in magnitude after preforming the projection followed by the back-projection operations [CEGL10]
- low-density objects are penalized by showing a quite high VMR value due to their very low mean value.

These two aspects are illustrated in Figure 3.35(d): in the image periphery where air voxels are reconstructed, the computed VMRs are too large, because the CT voxels there have small values (≈ 0). Moreover, the VMR map is too smooth to be used for localization of potential artifact voxels. Both issues can be solved by filtering the input projections in a way that the input signal is amplified proportional to the spatial frequency and the projection intensity (see Figure 3.35(b)). For this task the X-ray images are transformed to frequency domain, filtered with ramp-filter, then transformed to the spatial domain again (similar to the pre-filtering applied in the FDK algorithm). In order to compensate for higher frequencies, i.e. for noise, the basic ramp-filter is multiplied with Hamming window.

It is to be noted that the filtered projections $\widehat{P}_\theta$ may also have negative values, which may falsify the VMR values. Therefore the global minimal value ξ over all filtered projection data (all pixels in all X-ray images)

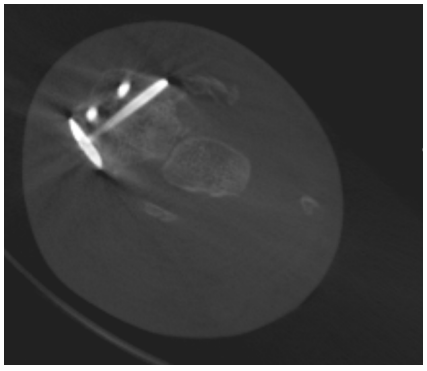
is added to all back-projection values (pixels). The Equation 3.19 is reformulated as

$$VMR(i) = \frac{\frac{1}{Q-1} \cdot \sum_{q=1}^{Q} \left(\widehat{P}_\theta(q_{i,\theta}) - \overline{\widehat{p}}_i \right)^2}{\overline{\widehat{p}}_i - \xi} \tag{3.20}$$

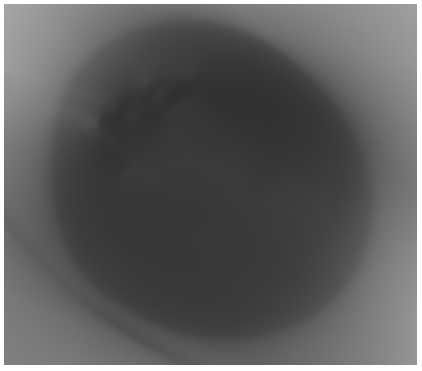
where $\overline{\widehat{p}}_i$ is the mean intensity of the pixels $\widehat{P}_\theta(q_{i,\theta})$ after applying the ramp-filter.

(a) Original X-ray image

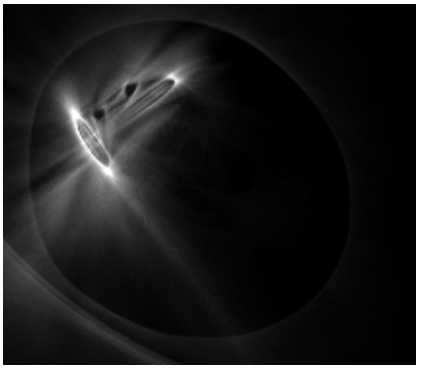
(b) Filtered X-ray image



(c) Representative CT slice exhibiting metal artifacts



(d) VMR-map computed based on unfiltered original X-ray images



(e) VMR-map computed based on filtered X-ray images

Figure 3.35.: An example illustrating the difference between VMR-maps computed based on the original X-ray images (a) and on filtered images (b). Unlike the VMR-map (d) calculated from (b), the artifact regions can be easily identified in the VMR-map (e) gained from (b). For comparison, the corresponding CT slice is shown in (c).

For illustration purposes three representative voxel types (a ”high-density/metal voxel”, an ”artifact voxel”, and an ”artifact-free voxel’) are shown in Figure 3.36. For each voxel type an example histogram of the corresponding projection values is plotted and the corresponding VMR is computed. An example showing the correlation between the artifacts levels and the proposed ratio is depicted in Figure 3.37.

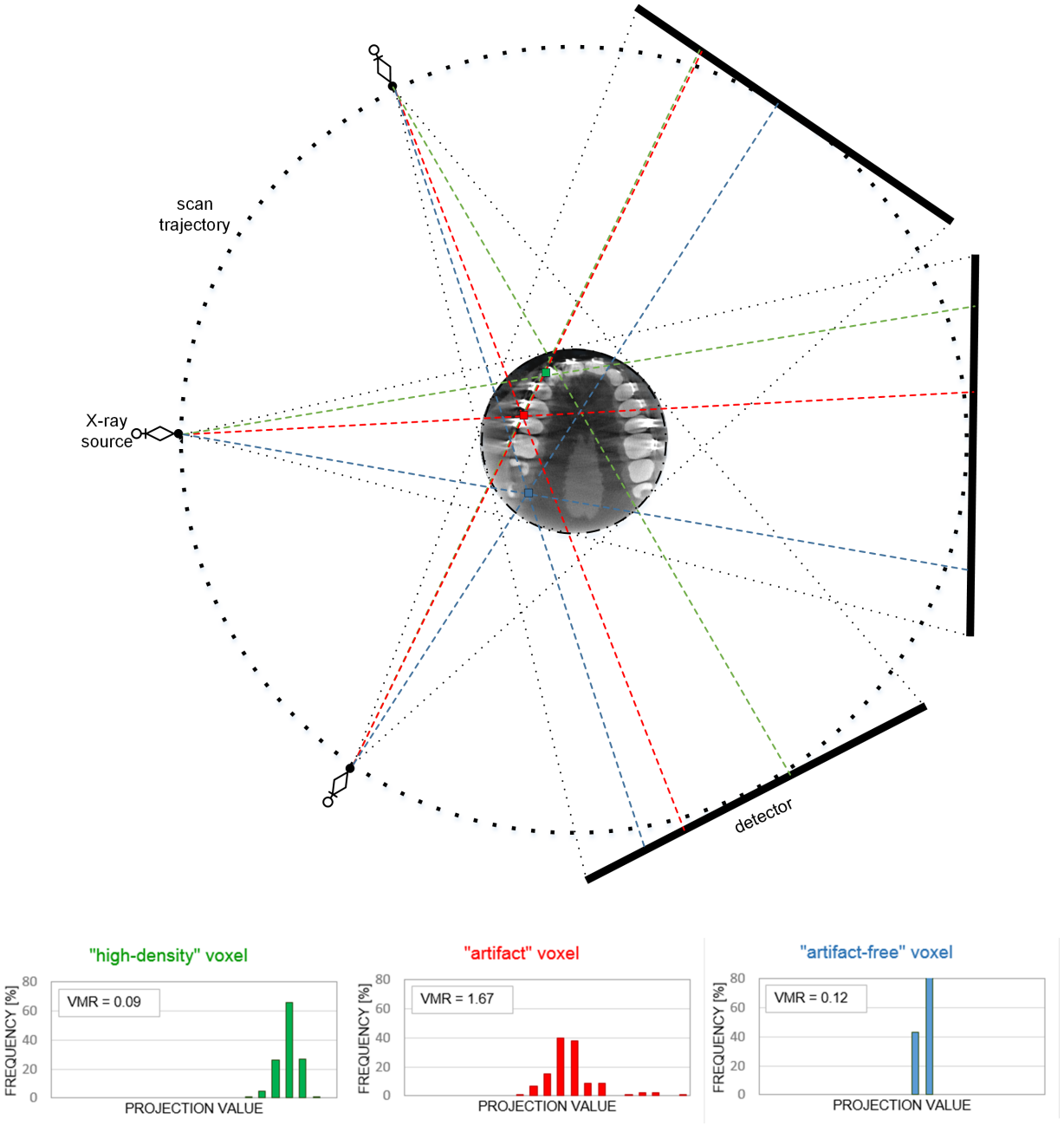
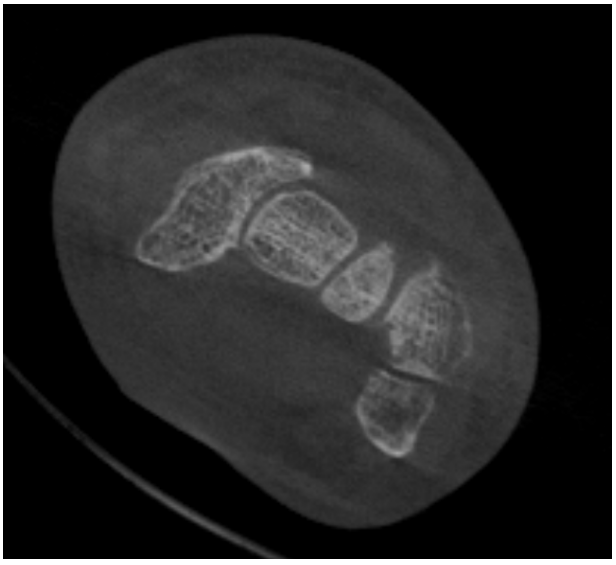
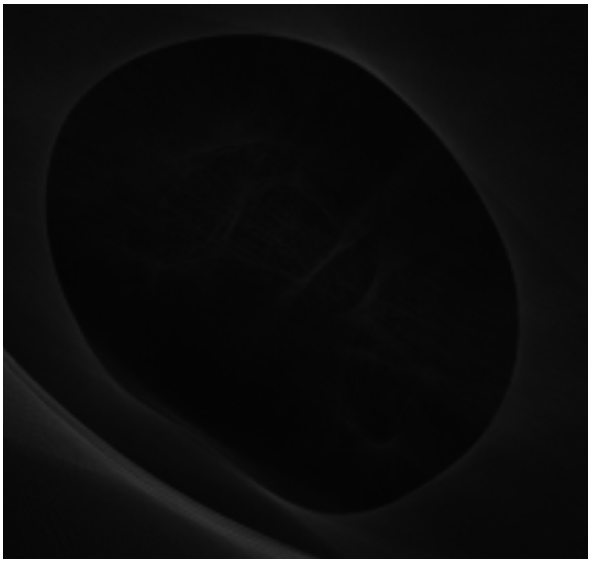


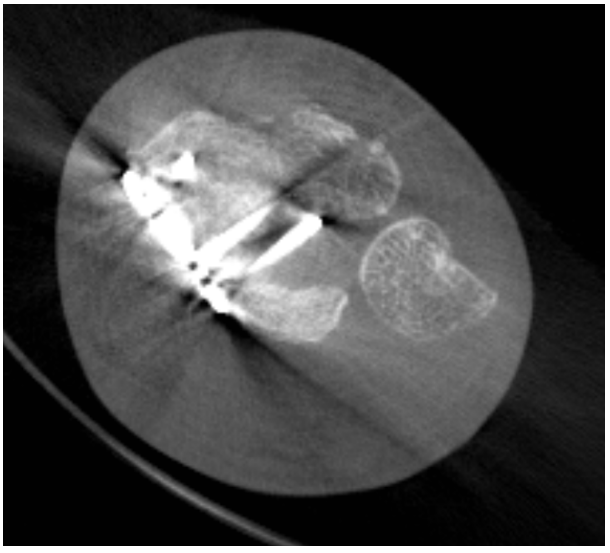
Figure 3.36.: Variance-to-mean ratios computed for three representative voxel types: a high-density voxel (green), an “artifact” voxel (red), and an “artifact-free” voxel” (blue). The corresponding histograms of the projection values, and the variance-to-mean ratios are plotted on the right side.



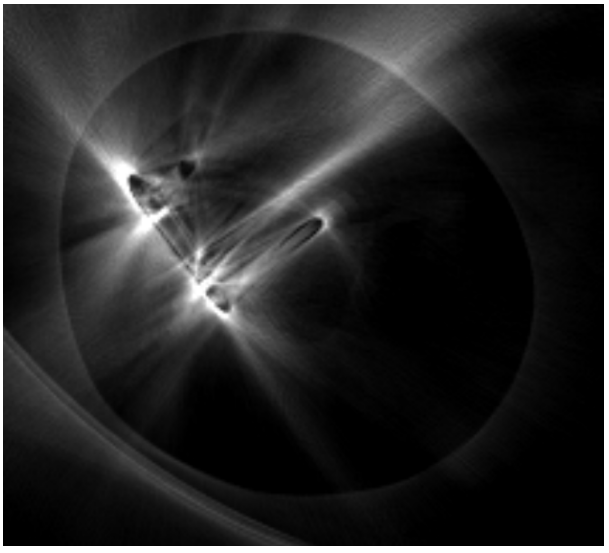
(a) Representative CT slice which is free of metal artifacts



(b) Variance-to-mean-ratio corresponding to the reconstructed slice (a)



(c) Representative CT slice which exhibits metal artifacts



(d) Variance-to-mean-ratio corresponding to the reconstructed slice (c)

Figure 3.37.: An example illustrating the correlation between the variance-to-mean ratios and the artifacts levels. A CT slice free of metal artifacts is shown in (a); the corresponding VMR-map (b) exhibits small values in the whole region. In contrast, the slice shown in (c) is affected with metal artifacts. In the corresponding VMR-map (d) the high-density voxels (metal parts) as well as artifact-free voxels have low VMRs. They appear as dark regions unlike artifact voxels, such as the voxels surrounding the "metal" voxels, which exhibit high VMRs and appear as bright regions.

Weighted similarity measure. In order to make the registration process robust against such artifacts (see Section 3.3.3), a weighted similarity measure is proposed. It aims to underweight the influence of the artifact regions on the registration result. Therefore, a spatially varying weighting function is developed to account for CT artifacts based on the variance-to-mean ratio computed for each volume voxel. As shown in Section 3.5.3, the VMR can be considered as a measure to estimate the probability that a given volume voxel belongs to an artifact region. As a consequence, voxels with high VMRs project on pixels which may be considered as potential artifact regions in the DRRs and should therefore have a reduced influence on the registration results. In contrast, unobscured as well as high-density voxels, i.e. voxels of low VMRs, are usually not affected by artifacts. Thus, the corresponding pixels are not penalized by the similarity measure.

max voxel value along $\vec{r}$	max VMR along a projection ray $\vec{r}$: low	max VMR along a projection ray $\vec{r}$: high
low	artifact-free pixel (air/soft tissue)	artifact pixel
high	artifact-free pixel (metal)	artifact pixel on which metal implants are projected

(a) Projection space (DRRs)

Voxel value	VMR: low	VMR: high
low	air or soft tissue voxel	CBCT artifact (dark streaks)
high	metal voxel	CBCT artifact (bright streaks)

(b) Voxel space (CBCT)

Figure 3.38.: Possible configurations of max VMR (along a projection ray $\vec{r}$) and density values, and their meaning for the registration and the CBCT reconstruction. The green color indicates in (a) pixels which should be considered in the similarity measure, in contrast to the red color which refers to pixels which should be ignored. In (b) the green color indicates the artifact-free voxels, whereas the red color indicates artifact voxels.

During the generation of the DRRs, the maximum VMR along each projection ray $\vec{r}$ is calculated. If the found VMR value is high (>1) and no high-density voxels are traversed by the said ray $\vec{r}$, the corresponding pixel is expected to contain reconstruction artifacts. In contrast, if the ray $\vec{r}$ traverses artifact voxels (high VMRs) as well as "metal" voxels (high attenuation values and low VMRs), the projection location of $\vec{r}$ is considered as artifact-free, since the said ray is always highly attenuated by the metal parts, weather artifact voxels are traversed by $\vec{r}$ or not. The four possible configurations (VMR, density value) and their meaning for the registration are summarized in Figure 3.38(a). The green color indicates in the pixels which should be considered in the similarity measure, whereas the red color refers to the pixels which should be ignored.

Let $\max_{\vec{r}}(V)$ and $\max_{\vec{r}}(VMR)$ be the maximum voxel value and the maximum VMR along $\vec{r}$, respectively. The weighting factor $w(q)$ computed for each projection pixel q is then defined as

$$w(q) = 1 - \underbrace{\frac{1}{1 + \exp[-\gamma \cdot (\max_{\vec{r}}(VMR) - 1)]}}_{\text{A}} \cdot \underbrace{\frac{1}{1 + \exp[\gamma \cdot (\frac{\max_{\vec{r}}(V)}{thres_M} - 1)]}}_{\text{B}} \tag{3.21}$$

where γ is a tuning parameter and $thres_M$ is a threshold over which a voxel is considered as a high-density (metal) voxel. Typical values for $thres_M$ and γ are 1000 and 10, respectively. The weighting factor $w(q)$ is in the range of $]0.0, 1.0[$: the influence of the corresponding pixel on the similarity measure is proportional to the attributed weight. As shown in Figure 3.39, the term A (left axis) under-weights pixels corresponding

to voxels with high VMRs (>1.0). On the other hand, if the projection ray $\vec{r}$ crosses metal voxels, the corresponding projection pixel should be considered in the similarity measure. In this case, the term B (right axis) compensates the effect of the term A, so that the corresponding weight tends again to 1. An example illustrating the projected VMR-map created as described above is shown in Figure 3.40.

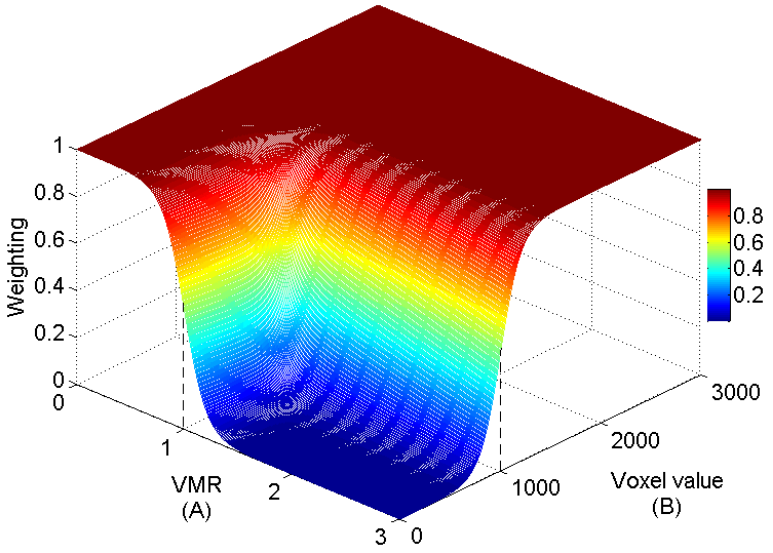
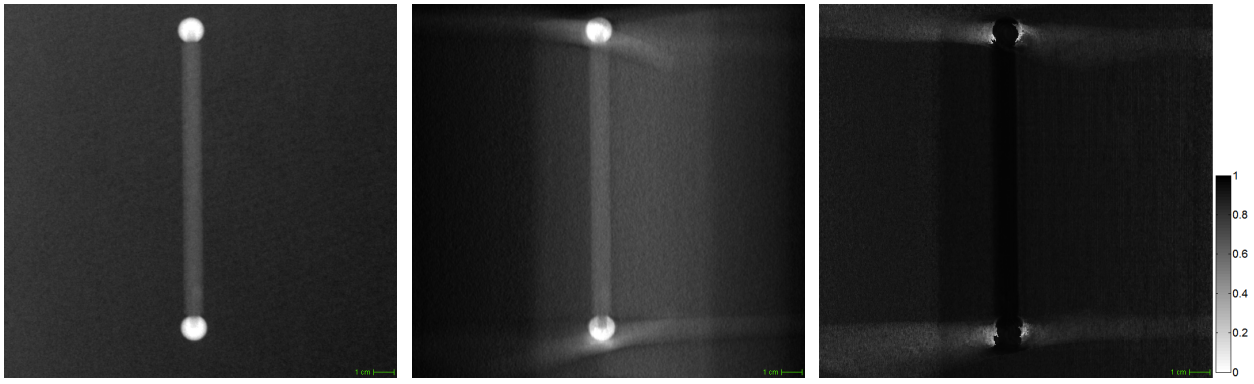


Figure 3.39.: VMR-based weighting function ($thres_M = 1000, \gamma = 10$).



(a) Acquired X-ray image (b) DRR (c) Weighting map

Figure 3.40.: A representative X-ray image acquired with the investigated C-arm CBCT system (a). Corresponding digitally reconstructed radiograph (DRR) and projected VMR-map are shown in (b) and (c), respectively.

In the following three weighted metrics are proposed and evaluated in the next section in term of registration accuracy and computational efficiency:

1. Weighted Sum of Squared Differences (WSSD):

$$WSSD = \frac{1}{N_O} \sum_{q \in O} w(q) \cdot (P_\theta(q) - (\widetilde{P}_\theta(T(q))) \tag{3.22}$$

where N_O is the number of pixels in the overlap domain O and T is the estimated transformation mapping the pixels with coordinates (q) of P_θ (originally acquired X-ray image) to $\widetilde{P}_\theta$ (simulated projection).

2. Weighted Normalized Cross Correlation (WNCC):

$$WNCC = \frac{\sum_{q\in O} w(q)\cdot(P_\theta(q)-\overline{P}_\theta)\cdot(\widetilde{P}_\theta(T(q))-\overline{\widetilde{P}}_\theta)}{\sqrt{\sum_{q\in O} w(q)^2\cdot(P_\theta(q)-\overline{P}_\theta)^2\cdot(\widetilde{P}_\theta(T(q))-\overline{\widetilde{P}}_\theta)^2}} \tag{3.23}$$

where $\overline{P}_\theta$ and $\overline{\widetilde{P}}_\theta$ are the average pixel intensities of the input projection image (P_θ) and of the DRR ($\widetilde{P}_\theta$), respectively.

3. Weighted Mutual Information (WMI):

$$WMI = \sum_{P_\theta}\sum_{\widetilde{P}_\theta} w(q)\cdot\mathcal{P}(P_\theta,\widetilde{P}_\theta)\log\frac{\mathcal{P}(P_\theta,\widetilde{P}_\theta)}{\mathcal{P}(P_\theta)\mathcal{P}(\widetilde{P}_\theta)} \tag{3.24}$$

where $\mathcal{P}(P_\theta)$ and $\mathcal{P}(\widetilde{P}_\theta)$ are the marginal distributions of the respective images P_θ and $\widetilde{P}_\theta$. The term $\mathcal{P}(P_\theta,\widetilde{P}_\theta)$ is corresponding joint distribution.

3.5.4. Evaluation of calibration accuracy

In order to verify the efficiency of both, the calibration algorithm and the proposed similarity measures, experiments have been performed using simulated as well as real data acquired with the investigated C-arm CBCT system operating in 3D scan mode (see Section 3.2.1). The experiment parameters are summarized in Table 3.3.

	Simulated data	Real data
Source-detector distance (in mm)	1064	1064
Pixel width/height (in mm)	0.283	0.283
Detector dimension (in pixels)	1024×1024	1024×1024
Angular sampling (in °)	2	1.5
Number of projections	180	111

Table 3.3.: Acquisition geometry parameters

Simulated CBCT data. The evaluation of the proposed calibration method was performed on two stages:

1. proof of concept: the validity of the proposed calibration process was checked based on simulated data without CT artifacts as described below. We refer to this case as test suite (A).
2. adequacy of the proposed weighted similarity metrics in case of metal-artifacts. We refer to this case as test suite (B).

Both of these test suites were performed following the same scheme described below. However, in the test suite (A) we employed only artifact-free images (ideal detector linear response), whereas in the second suite (B) more realistic data sets were created, namely CT and DRRs affected by metal-artifacts. In the next, we described how we proceeded to create the data sets for both tests, and then we list the achieved results.

The calibration accuracy was quantitatively evaluated by simulating several phantoms and ideal orbits. For each phantom a CT volume CT_{ref} was created and then ideal DRRs, called here "ground truth images – GTI", were generated from CT_{ref} based on the known ideal geometric projection parameters. In the test suite (A) ideal DRRs were generated by simulating linear response for the detector. Consequently, the CBCTs reconstructed from these DRRs would not contain any metal-artifacts, unlike the DRRs created for the test suite (B):

- the simulated phantoms contained soft tissue as well as high-density objects
- the non-linear response of the detector was simulated while generating the DRRs (GTIs), and therefore the saturation effect as well as the high-attenuation of X-rays by metallic parts were considered. Consequently, CBCT volumes CT_{rec} reconstructed based on the GTIs contained metal-related artifacts
- any DRR which is generated based on CT_{rec} was *not* artifact-free, since CT_{rec} was affected by metal-related artifacts.

For each of the GTIs all of the nine reference geometric projection parameters (RPP) was corrupted one at a time with a Gaussian distributed error ($\mu = 0$ mm; $\sigma_{trans} = 0.3$ for translational parameters and $\sigma_{rot} = 0.03$ for rotational parameters) resulting in erroneous geometric projection parameters (EPP). The GTIs and the corresponding modified geometry data EPPs has been then provided as input for the calibration process, which should recover the original, correct RPPs. As metric for success we considered the target registration error (TRE) between the GTIs and the corresponding DRRs after successful fine-tuning, i.e. the resulting projection error on the detector plan, which is very significant for the 3D reconstruction quality. Accordingly, for any given gantry angle θ the calibration error ϵ_θ is defined as

$$\epsilon_\theta = \max_{\vec{x}}\left(\left\| H_\theta \cdot \vec{x} - \widehat{H}_{\theta,x_s} \cdot \vec{x} \right\|_2\right) \tag{3.25}$$

where $\vec{x} = (x, y, z, 1)$ are the voxel coordinates, and ϵ_θ is the maximum deviation calculated at the gantry angle θ by iterating over all volume voxels $\vec{x}$.

Test suite (A): Proof of concept. Here, the DRRs are defined to be identical to the ground truth images (GTIs). In this case, a simple metric such as the sum of squared difference (SDD) should be sufficient for good registration results due to the consistency of the pixels values in both sets (DRRs and GTIs). As shown in Figure 3.41, the correct geometric parameters could be recovered with a mean uncertainty of 0.08 pixels in the projection domain (on the X-ray image). The remaining error is mainly caused by the creation of the DRRs due interpolation errors.

The main challenge in real application is the presence of CT artifacts. As it will be shown below, the SDD metric is not anymore capable to achieve good calibration results. In the next test suite the focus is set on the influence of the similarity metric on the calibration accuracy in case of images affected by metal artifacts.

Test suite (B): Adequacy of proposed weighted similarity metrics. In order to analyze the influence of the similarity measure itself as well as the proposed weighting scheme on the fine-tuning accuracy, tests were repeated for different similarity metrics, applied with and without weighting respectively. As depicted in Figure 3.42:

- The sum of squared differences (SSD) shows large deviations due to the large intensity differences expected in the artifacts regions, leading to spread the SDD values. It is possible to limit the influence of huge differences between the intensities by using the sum of absolute difference instead of SSD, but this would decrease the sensitivity of the metric against misalignments.
- In case of mutual information (MI) the intensities in the artifacts regions lead to change the spreading of the probability distribution of registered images. Thus, the registration error may significantly increase.
- The normalized cross correlation (NCC) provided slightly better results in comparison to the SSD.

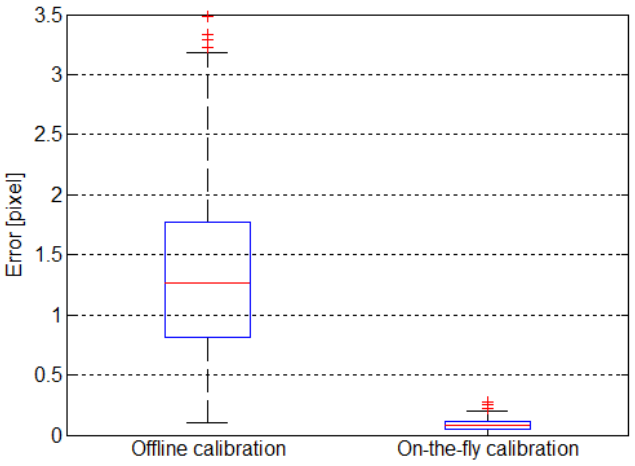


Figure 3.41.: Calibration uncertainty in case of artifact-free CT and DRRs.

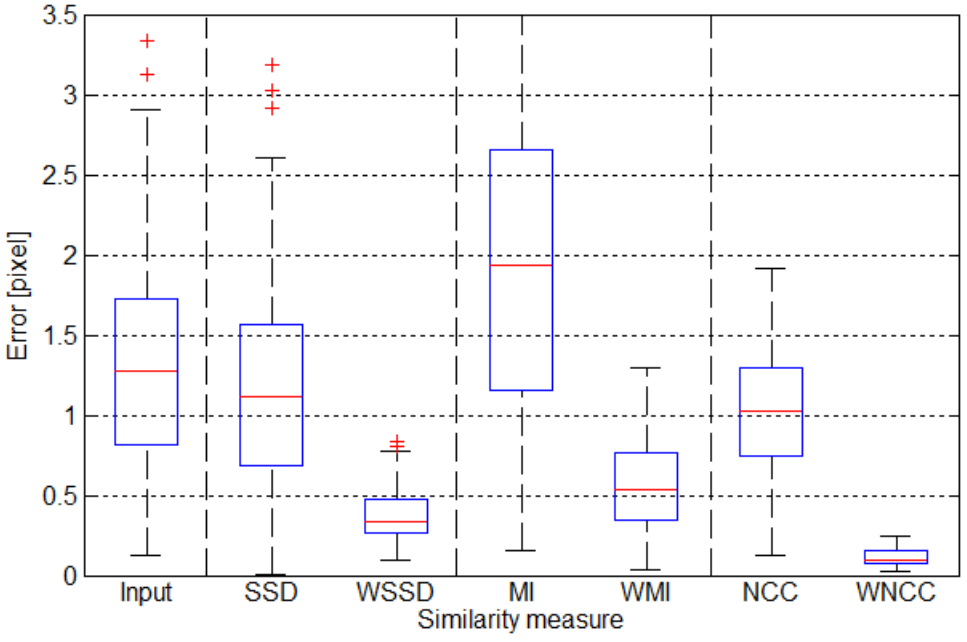


Figure 3.42.: Calibration uncertainty for several similarity measures: sum of square differences (SDD), weighted SSD (WSSD), mutual information (MI), weighted MI (WMI), normalized cross correlation (NCC) and weighted NCC (WNCC).

By employing the proposed weighting function, the calibration uncertainty is significantly reduced for all metrics because the registration results are less affected by the artifacts regions. The mean error is thereby reduced for all similarity methods to sub-pixel range.

Best results are achieved with the weighted NCC metric. As shown in Figure 3.43, the mean error is reduced from 1.38 pixels to 0.11 pixels, and the maximum error from 3.53 pixels to 0.24 pixels, i.e. an order of magnitude smaller. Therefore this similarity method is employed for the proposed on-the-fly calibration. As depicted in Figures 3.43, a sub-pixel accuracy is achieved in the field of view within 2 iterations. However, the elimination of all outliers requires up to five iterations (see Figures 3.44 and 3.44). Such improvements is beneficial for the 3D reconstruction quality, as visually illustrated in Figure 3.45.

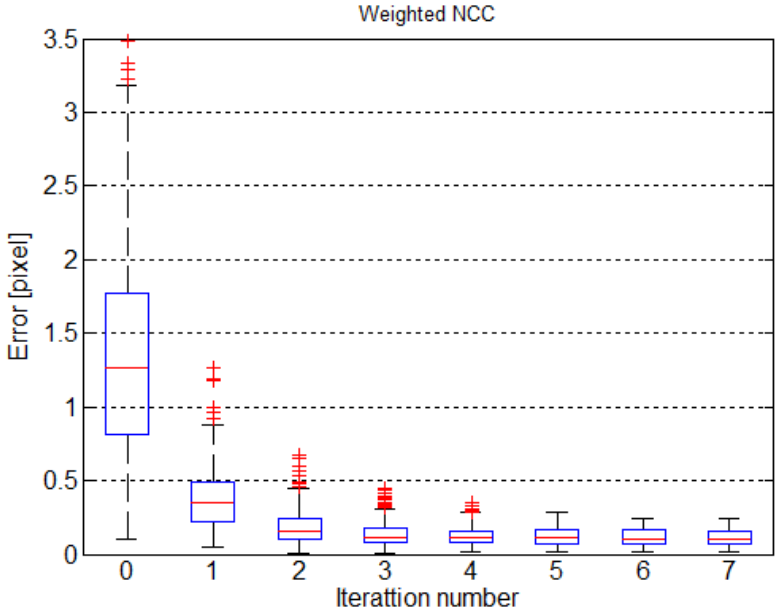


Figure 3.43.: Calibration uncertainty as function of iteration number. The pre-calibrated data (output of the offline calibration) corresponds to the iteration number 0.

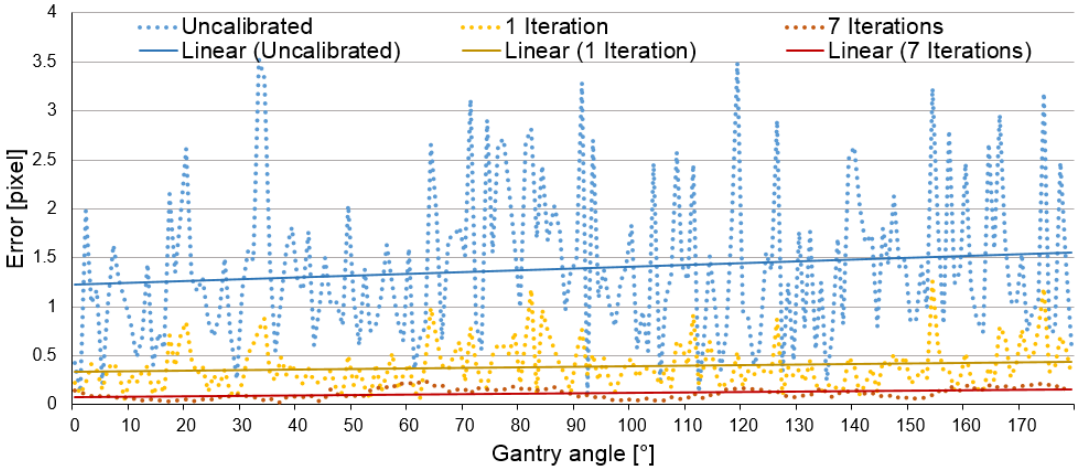
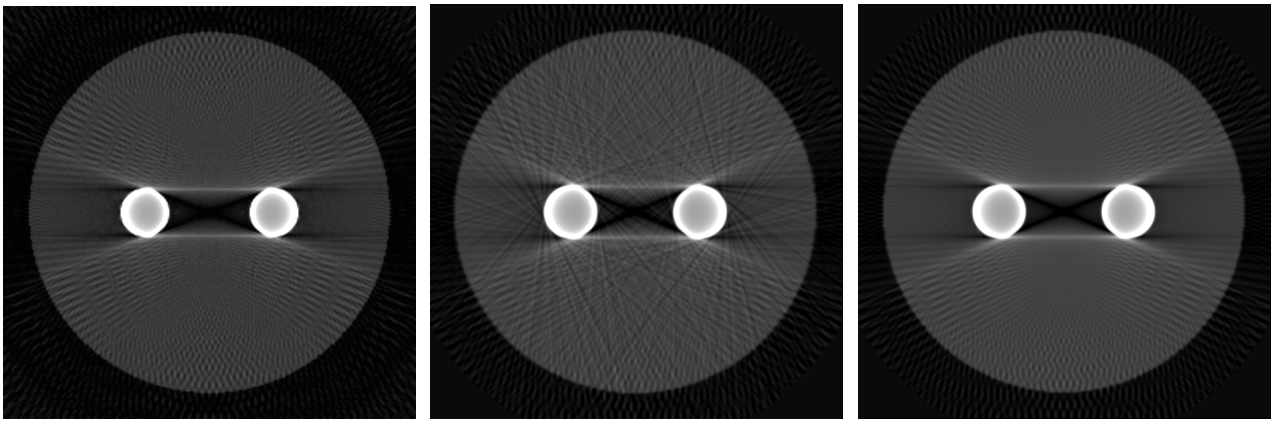


Figure 3.44.: Uncertainty of the on-the-fly calibration given in pixel units. The employed similarity measure is the proposed weighted normalized cross correlation metric (WNCC).

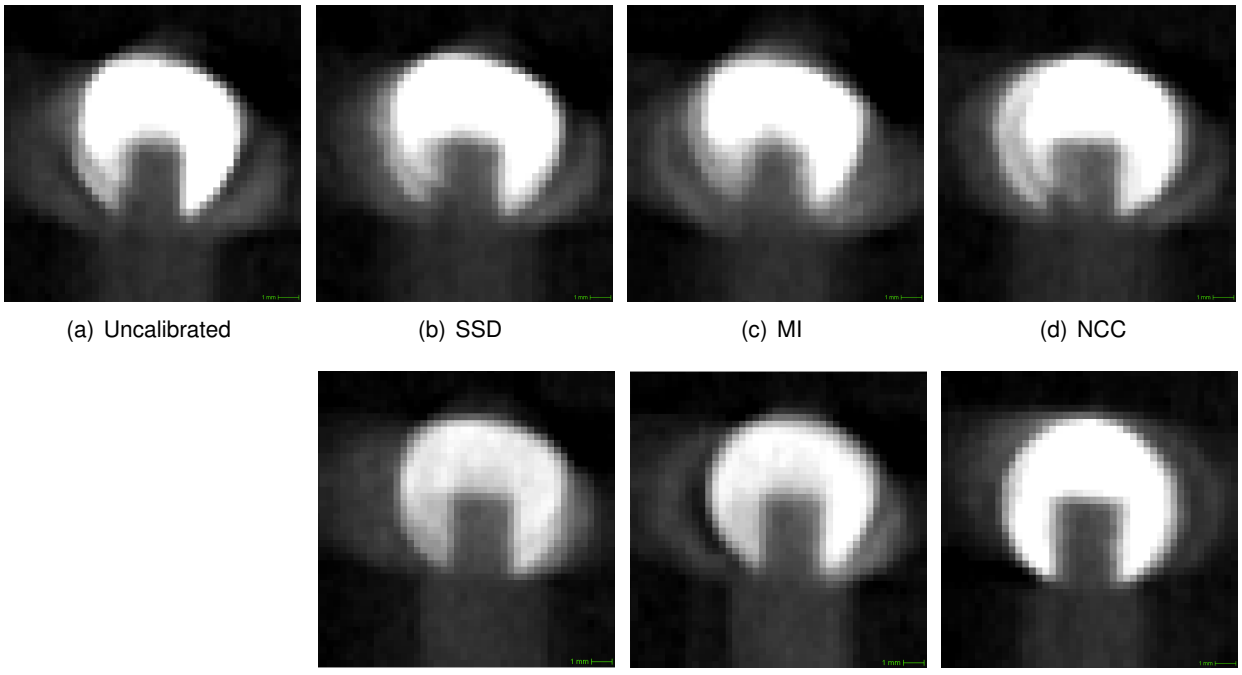


(a) Reference image (b) Pre-calibration (c) Proposed on-the-fly calibration

Figure 3.45.: Example illustrating the improvement in the image quality after applying the proposed on-the-fly (online) calibration. The reference image corresponds to the reconstruction preformed based on the correct geometric projection parameters used to simulate the X-ray projections. The three volumes are reconstructed with the FDK algorithm.

Real CBCT data. In case of real data, the ground truth is unknown. Thus a direct quantitative evaluation is not possible, however an indirect evaluation can be done by visually inspecting the reconstruction quality. A double-ball-bar phantom was scanned for this task. As shown in Fig. 3.46 and Fig. 3.47, the spherical shape of the balls is lost when only pre-calibration parameters are used for the 3D reconstruction. After performing the proposed on-the-fly calibration in combination with the VMR-weighted NCC, the balls are correctly reconstructed as spheres (see Figure 3.46(g)). Even small details, pointed out with red arrow in Figure 3.47, which are completely lost due to geometrical misalignments are recovered.

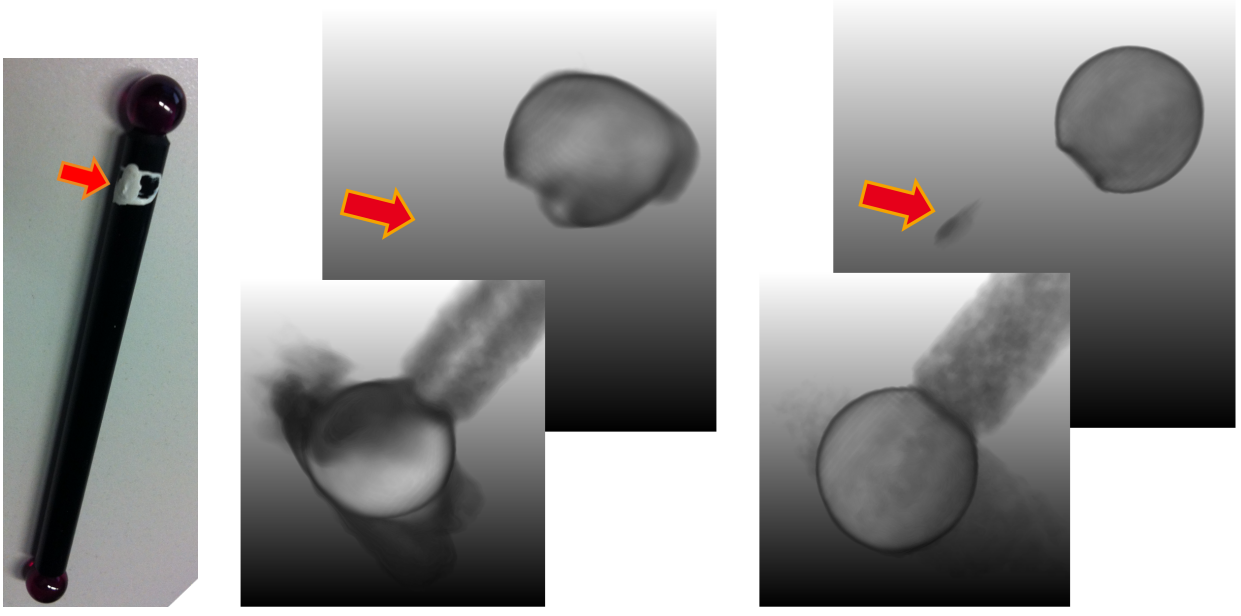
Computational complexity. The tests were performed on an Intel Core i5-4670 CPU with 3,40 GHz CPU and 16 GBytes RAM. The DRR generation was performed on a GeForce 7600 GTX graphics card. The size of the CT images and the X-ray images was $512\times512\times512$ and 1024×1024, respectively. Table 3.4 summarizes the processing time required to perform the no-the-fly calibration for single X-ray image. By reducing the geometric parameters to be estimated, the calibration process gets 18 times faster compared to the original formulation without parameters reduction. With the help of parameter reduction the registration process is reduced from 2D-3D registration to rather trivial 2D-2D registration. In this case, only one DRR image has to be generated. In contrast, the 2D-3D registration requires the generation of a new DRR in each iteration, which is very time consuming. Taking into account that the CBCT reconstruction is performed based on more than 110 X-ray images, it is obvious that only with the help of the parameter reduction it is possible to achieve the 3D reconstruction within 1min, which is acceptable for clinical purposes. For instance, the on-the-fly calibration parameter reduction requires less than 31s if it is palatalized over 8 threads. For comparison purpose, the same calibration takes 9min 53s if no parameter reduction is employed.



(a) Uncalibrated (b) SSD (c) MI (d) NCC

(e) Proposed WSSD (f) Proposed WMI (g) Proposed WNCC

Figure 3.46.: Representative slices illustrating the influence of the on-the-fly (online) calibration, combined with different t similarity measures, on the reconstruction quality.



(a) Phantom (b) Offline calibration (c) Proposed on-the-fly calibration

Figure 3.47.: Volume-rendering of 3D images reconstructed from X-ray images of a double-ball-bar phantom (a) before (b) and after (c) performing the proposed calibration method. The spherical shape of the ball is better reconstructed by employing the on-the-fly calibration (combined with WNCC). Even some small structures (red arrow) could be recovered.

	Without parameters reduction	With parameters reduction	Speed-up factor
min	13.41 s	0.68	19.72
max	14.44 s	0.80	18.04
mean	13.74 s	0.73	18.82
std	0.42	0.04	–

Table 3.4.: The execution times needed for the on-the-fly calibration of a single 2D X-ray image.

3.6. Iterative VMR-based reconstruction

The presence of high-density objects, such as metal implants, may cause metal artifacts in the reconstructed images. These artifacts have prominent streak-like appearance and balk-like shadows between high-density objects visible as so-called “background artifacts”. This thesis contributes a novel 3D reconstruction algorithm aiming to reduce the background artifacts and enhance the contrast in the soft tissue regions in case of moderate metal artifacts. An extension of the proposed algorithm to deal with heavy metal artifacts is presented in the next section.

For illustration purposes, a phantom made of a water cylinder containing two metal spheres is employed (see Figure 3.48(left)). The mentioned typical metal artifacts appear in the CBCT volume reconstructed with FDK algorithm as dark/bright streaks connecting the metal objects in the directions of highest attenuation, as shown in Figure 3.48(right). In case of real data these CT artifacts are usually even more excessive. They are mainly caused by non-linear detector response: in the detector regions shadowed by the high-density parts, a nonlinear detector response as well as an increase in noise are expected. This can result in streaking artifacts due the discrepancy between the modeled linear detector response and the physical signal acquisition process. In real CBCT data these artifacts are even more excessive due to scatter.

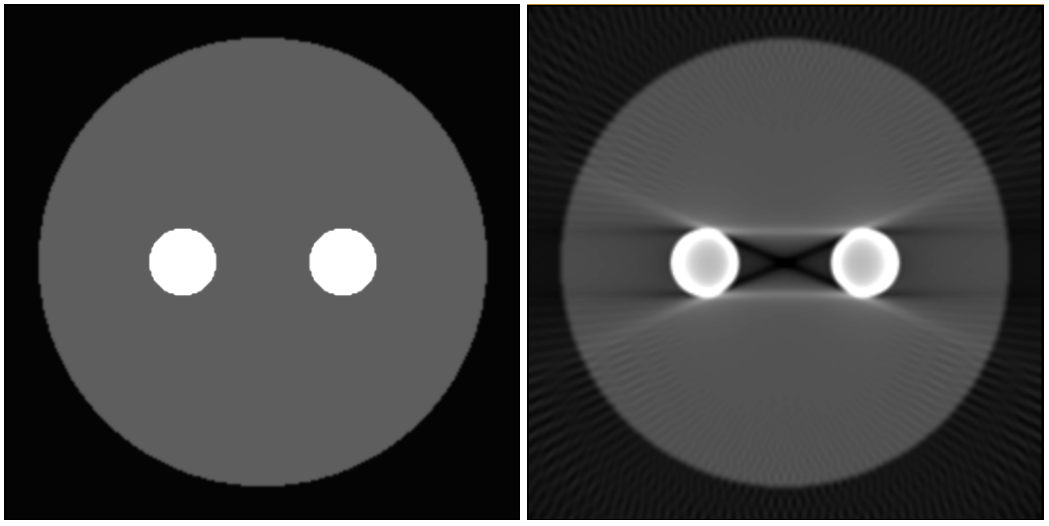


Figure 3.48.: A representative slice of a phantom made of a water sphere with two metal spheres inside (left) and the corresponding CBCT volume, reconstructed with FDK algorithm, which is affected by typical CT artifacts (right). The shadow and streaks are not limited to the regions around the spheres, but corrupt the whole image.

3.6.1. Concept

As shown in Section 3.3.3, the reconstruction problem can be formulated as a linear system $AV = P$, where $A = [a_{iq}]$ is the matrix mapping the volume voxels $V(i)$ to the projection pixels $P(q)$ with a_{iq} the contribution of the voxel i to the pixel q. Solving this problem by direct inversion technique is impractical. Thus, we follow an iterative scheme as described in section 3.3.3.

The main idea of the proposed method is the use of the VMR-map to perform a non-linear reconstruction which considers the probability that a given voxel is subject to be affected by CT artifacts. The process pipeline, shown in Figure 3.18, consists of the following steps:

1. Perform an initial CBCT reconstruction with FDK algorithm which provide good initialization with very low computational cost
2. Compute the variance-to-mean ratio for each voxel, as described in Section 3.5.3. This step serves to estimate the probability that a given voxel could be affected by CT artifacts.
3. Calculate a DRR for each input X-ray image based on the current reconstructed CT volume and the corresponding projection parameters
4. Compare the DRR to the input X-ray image and compute a correction term for each pixel
5. Rescale the correction term as described below in the back-projection formula
6. Back-project each correction term along the corresponding projection ray and update the voxels traversed by this ray
7. Iterate over the steps 2 to 6

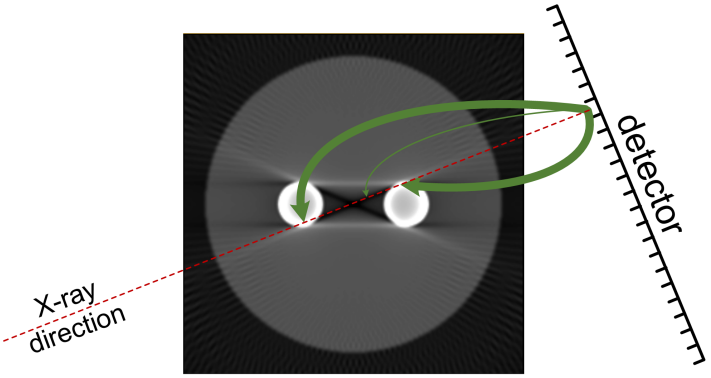
Back-projection formula. For each X-ray image, a DRR is generated based on the current CT volume V^k. The difference between the X-ray image and the DRR is employed to create a new estimate of V^k as follows:

1. for each pixel q of the current X-ray image, the correction term $\epsilon_q = P_\theta(q) - \widetilde{P}_\theta(q)$ is calculated
2. ϵ_q is multiplied with a weighting term α_i
3. the weighted correction value is back-projected along the projection ray $\vec{r}_q$.

As shown in Section 3.3.3, α_i is usually kept constant over all projection pixels or set proportional to the intersection length between the voxel of interest and the projection ray $\vec{r}_q$ [GBH70, KS01]. In both cases a linear distribution of the correction term over all voxels along $\vec{r}_q$ takes place. This would be sufficient if no metallic objects are scanned, since the detector would have an approximately linear response in the whole field-of-view. However, such a linearity is not fulfilled at pixels on which metal implants are projected (see Section 3.3.3). Consequently, the correction terms computed at metal pixels are typically not as accurate compared to the pixels with linear response. Now, if these inaccurate correction terms (computed at metal pixels) have big impact on the final value of a given soft tissue voxel, this voxel would be reconstructed as CT artifact.

In order to overcome this problem, the correction terms should be better distributed over the voxels along the projection ray. Accordingly, correction terms computed at metal pixels should be mainly concentrated on the metal voxels, not on the soft-tissue voxels, as illustrated by the width of the green lines in Figure 3.49.

This thesis contributes a new weighting scheme which aims to non-linearly distribute the correction terms along the corresponding projection rays in order to reduce CT artifacts: Assuming that the positions of



(a) VMR-based weighting in case of metal pixels

Figure 3.49.: Illustration of non-linear weighting of the correction terms computed at metal pixels in order to concentrate them on the metal voxels. The width of green arrows is proportional to the correction term values.

metal pixels as well as metal voxels are known, the reconstruction kernel should concentrate the correction terms computed at metal pixels to the metal voxels. However, the positions of metal voxels are unknown at the beginning of the reconstruction process. Even if the CT volume is already initialized e.g with the FDK algorithm, the segmentation of metal voxels would be inaccurate, since this initial CT volume contains metal-artifacts which may appear similar to metal voxels and would therefore be wrongly segmented. In order to solve this problem we propose to make use of the VMR-map (see Section 3.5.3) which assigns a probability value to each voxel for being artifact-free or not. Typically, the value of voxels which have high VMRs are erroneously reconstructed if they are aligned with high density voxels. In such a case it is desirable to underweight the contribution of the metal pixels to these voxels. Therefore, the correction term to be back-projected is weighted by a VMR-term w_{VMR} which is defined as

$$w_{VMR}(VMR) = \frac{1}{1 + \exp(8 \cdot (VMR - 1))} \tag{3.26}$$

The term $w_{VMR}(VMR)$ is plotted in Figure 3.50 (blue line).

As illustrated in Figure 3.49, the proposed VMR-based weighting penalizes the artifact voxels over all projections, even in case where these artifact voxels are not obscured by metal objects on the X-ray images. Such a behavior would result in under-estimating the values of the artifact voxels, since they do not obtain sufficient (positive) contributions from the correct projections. In order to not over-penalize these voxels with high VMRs, the following two modifications are proposed:

- The VMR-based weighting term becomes less significant with the increasing iteration number k as depicted in Figure 3.50. Therefore, the equation 3.26 is reformulated as

$$w_{VMR}(VMR) = \frac{1}{1 + \exp((\frac{2}{k})^3 (VMR - 1))} \tag{3.27}$$

  However, this modification should not influence the main idea of focusing the correction term computed at metal pixels on the appropriate metal voxels. Therefore, the next modification is also proposed.
- Metal voxels are characterized by their high densities and low VMRs (see Figure 3.38(b)). This information can not be employed at the beginning of the reconstruction process, since all voxel values

are set to zero. However, after back-projecting 10-20 X-ray images, the values of metal voxels which increase quickly is high enough to be recognized. Now, if a soft tissue voxel (low density) is aligned with a metal voxel (high density), the contribution of metal pixel can be concentrated on metal voxels by making the correction term proportional to the current voxel value. Accordingly, the correction terms are weighted proportional to the voxel intensity V_i and normalized by dividing it through the maximum voxel intensity along the projection ray $\vec{r}_q$, i.e. proportional to $(1 + V(i))/(1 + \max_{\vec{r}_q}(V(j)))$. If metal voxels are crossed by the projection ray $\vec{r}_q$, the maximum value along this ray would be high. Consequently, the contribution to the soft tissue voxels, which have low values, would be small, unlike the contribution to metal voxels. In contrast, if no metal voxels lie on $\vec{r}_q$, the maximum value along this ray would be relatively small, so that the soft tissue voxels would not be penalized.

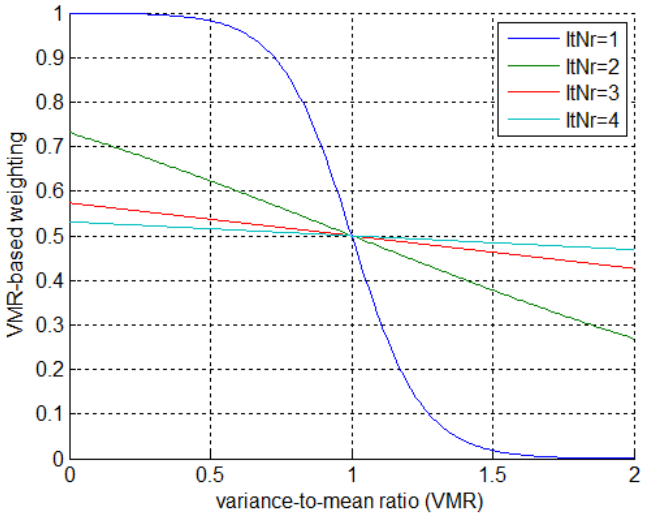


Figure 3.50.: VMR-based weighting of the correction term to be back-projected. The influence of the VMR decreases with the increasing iteration number.

Lastly, the back-projection term is also weighted proportional to the intersection length $l_{q,i}$ of the ray $\vec{r}_q$ with the voxel of interest at position i. A normalization is performed by dividing $l_{q,i}$ by the total ray length l_q.

Considering the proposed weighting functions, applied to the correction terms based on the characteristic of voxels crossed by the ray $\vec{r}_q$, the formula for the iterative reconstruction is formulated as follows

$$V^{k+1}(i) = V^k(i) + \lambda \cdot \frac{l_{q,i}}{l_q} \cdot \frac{1 + V^k(i)}{1 + \max_{\vec{r}_q}(V^k(j))} \cdot w_{VMR}(VMR(i)) \cdot (P(q) - \widetilde{P}(q)) \tag{3.28}$$

where λ is a relaxation factor which serves to control the convergence of the method: small λ values results in slow convergence, unlike large λ values. However, in the later case may cause divergent high frequency structures which develop during the iterations. In our experience, the λ value should be in the range $[0.5, 2.0]$. For a physically correct solution, negative values are not accepted and are rather set to zero.

3.6.2. Evaluation of iterative VMR-based reconstruction

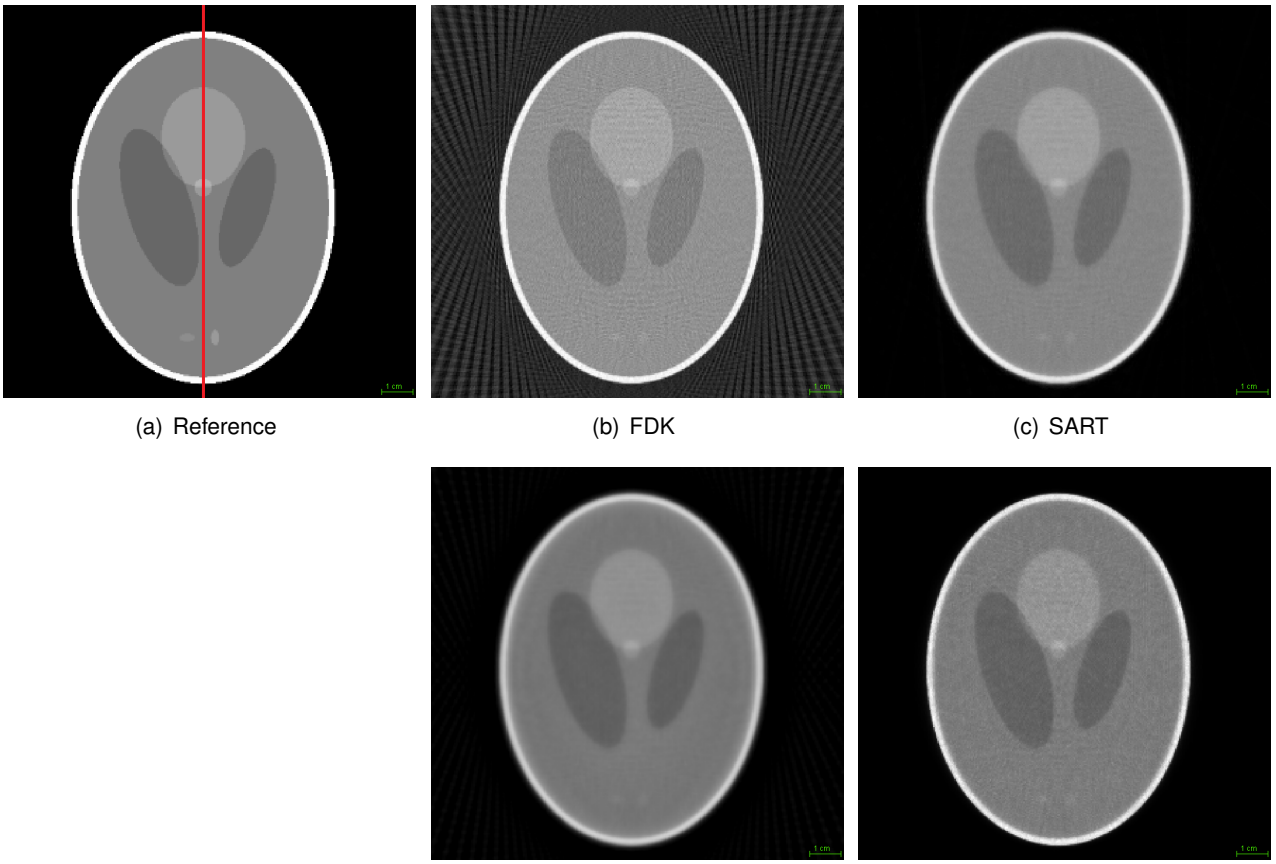
In this section quantitative and qualitative evaluations of the proposed reconstruction algorithm are performed. For comparison purposes several conventional reconstruction methods are investigated, namely the Feldkamp algorithm (FDK), the simultaneous algebraic reconstruction Technique (SART) and the iterative total variation based reconstruction (TV).

Only simulated data sets can deliver perfect ground truth for comparison. However, due the discrepancy between the imaging model and the physical scanning situation, additional artifacts prevalent in clinical situations are not reproducible in simulation-based studies. Therefore, tests were performed based on simulated as well as real data:

- computational phantoms: detailed computer models of a double-spheres phantom (see Figure 3.48) and the Shepp-Logan phantom (a model for the human head) [SL74] were created. From these phantoms appropriate CT volumes CT_{ref} were generated. These CTs were defined as ground-truth and based on them DRRs were simulated with geometrical parameters similar to those of the investigated C-arm CBCT system (see Table 3.3).
- CT volumes: a cadaver temporal bone was scanned with a SOMATOM Definition AS+ scanner in high-resolution mode. The reconstructed CT volume was considered as a ground truth and served to generate DRRs. The purpose here was to generate data sets which are more appropriate to study the reconstruction artifacts in the case of skull-base CTs. It is to be noted that the gantry angle was set to 360° in this simulation study, since the influence of the reduced scan angle typical for C-arm systems is out of focus in this thesis.
- real X-ray projection: data of a step-cylinder, a double-ball-bar, a head and a hand phantoms are acquired with the mentioned mobile C-arm CBCT system. In this case no reference CT (ground-truth) was available, thus the evaluation was limited to visual inspection.

Simulated Data. An axial image of the Shepp-Logan phantom, reconstructed based on projections simulated according to the geometry parameters listed in Table 3.3, is shown in Figure 3.51. It can be seen that the proposed approach (iVar) clearly outperforms the FDK algorithm. The TV algorithm tends to reconstruct relatively smoothy images. Compared to the SART algorithm, the proposed method (iVar) shows better contrast with less artifacts. This impression correlates with the quantitative evaluation plotted in form of intensity profile along the center-line (red line) shown in Figure 3.52. Similar results are obtained in case of the double-spheres phantom as depicted in Figure 3.53. The streak artifacts are clearly attenuated using the proposed iVar method while maintaining high contrast. The background artifacts are also strongly decreased in comparison to other reconstruction techniques.

CT Data. The temporal bone phantom has fine anatomic structures in the inner ear region and demands therefore high spatial resolution. As depicted in Figure 3.55, the background noise is clearly reduced using the proposed iVar algorithm. Representative axial, sagittal and coronal images of the different reconstructions of the middle ear are shown in Figure 3.56. The fine structures are also better reconstructed with the proposed iVar method. In contrast, the other algorithms, especially the TV approach, tend to fuse the fine edges.



(a) Reference (b) FDK (c) SART

(d) TV (e) Proposed (iVar)

Figure 3.51.: Reconstruction of the Shepp-Logan phantom with different reconstruction methods.

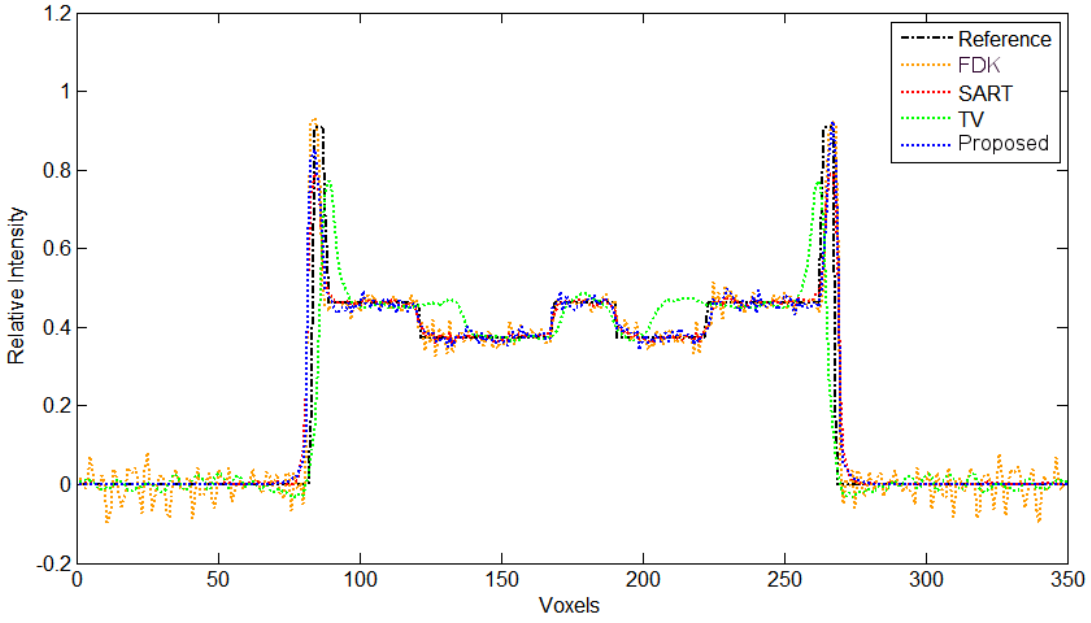
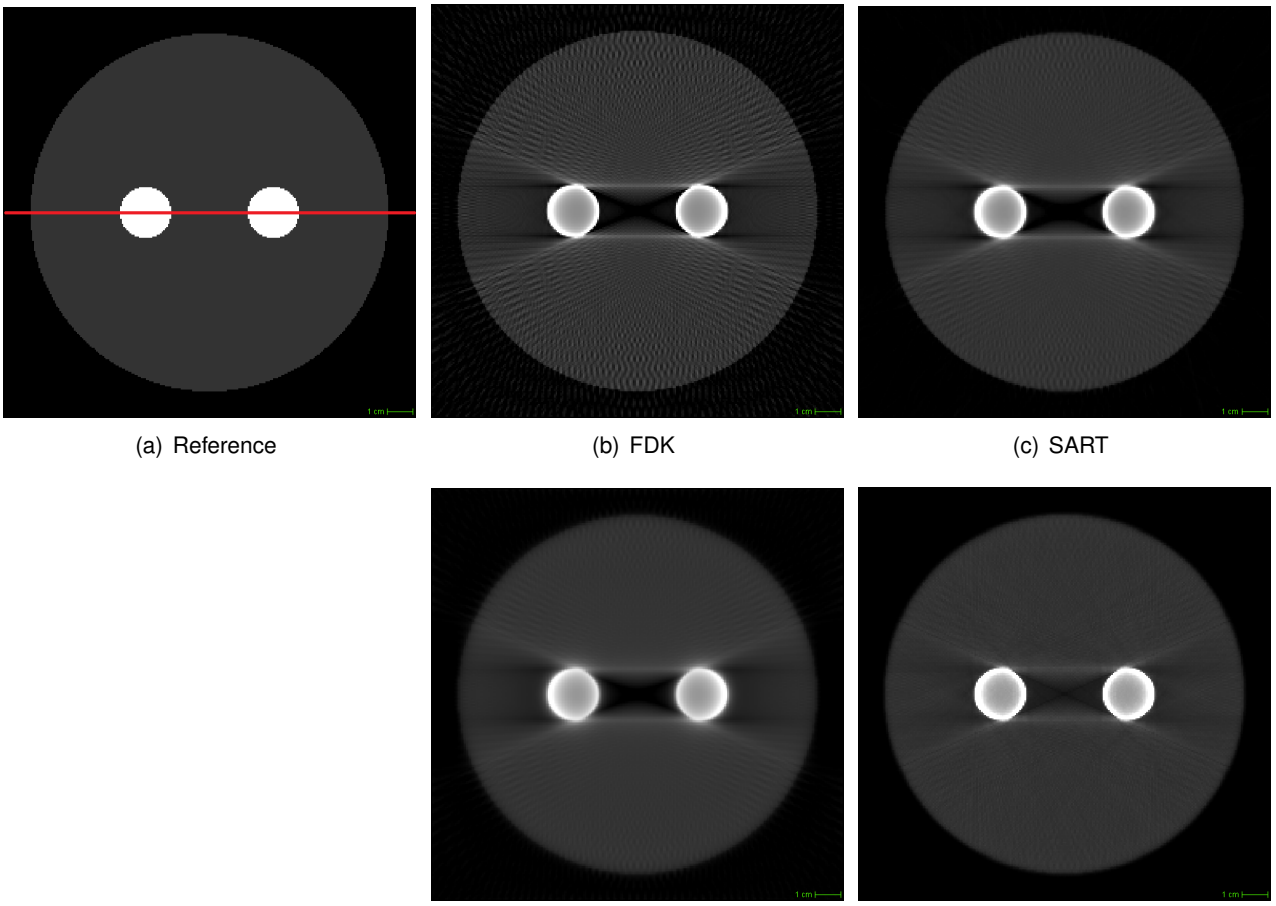


Figure 3.52.: Intensity profile of the reference and reconstructed volumes along the red line shown in Figure 3.51(a).



(a) Reference (b) FDK (c) SART

(d) TV (e) Proposed

Figure 3.53.: Reconstruction of the two-spheres phantom with different reconstruction methods.

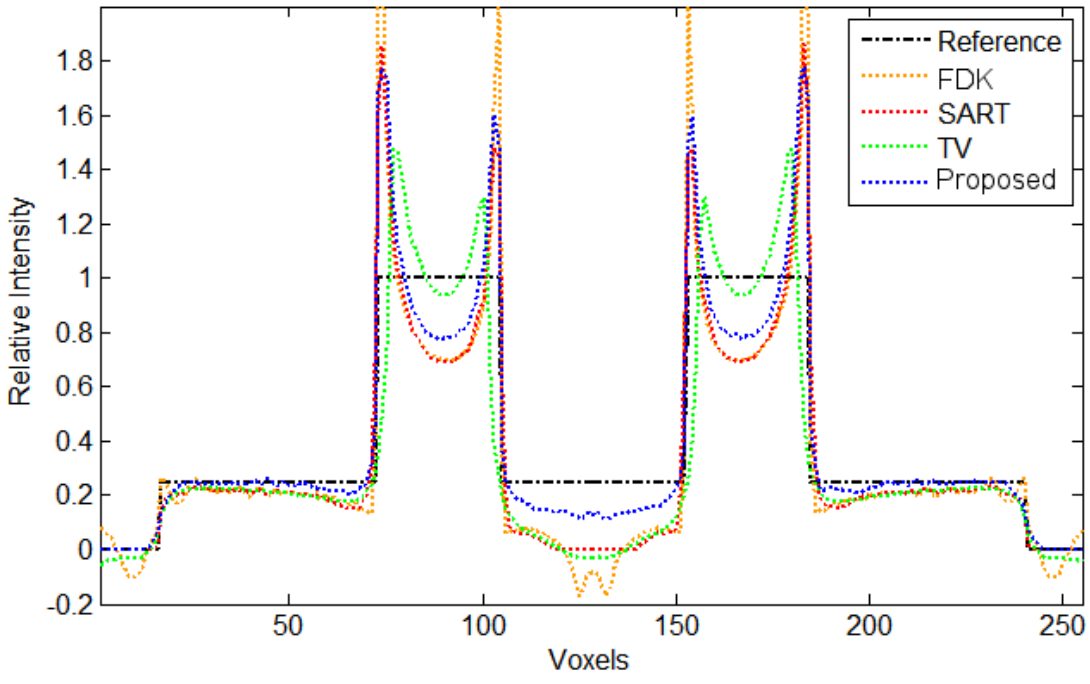
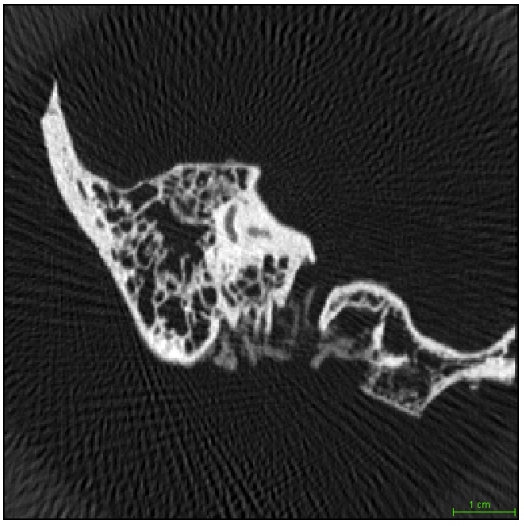
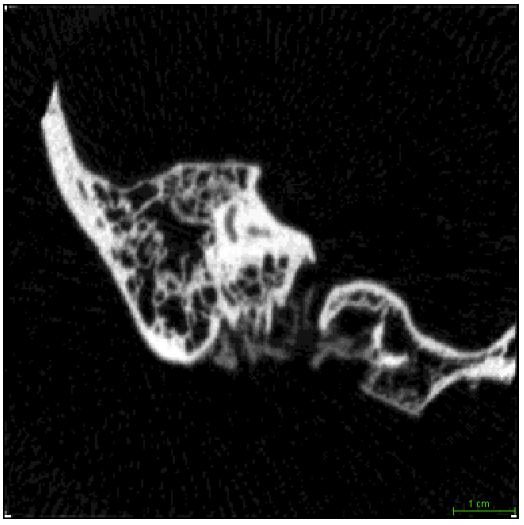


Figure 3.54.: Intensity profile of the reference and reconstructed volumes along the red line shown in Figure 3.53(a).

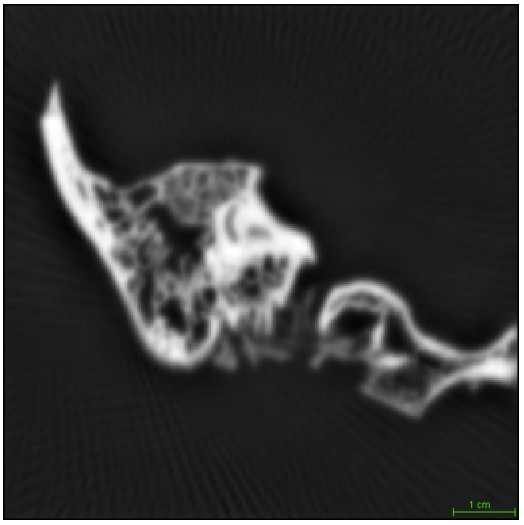
(a) Reference



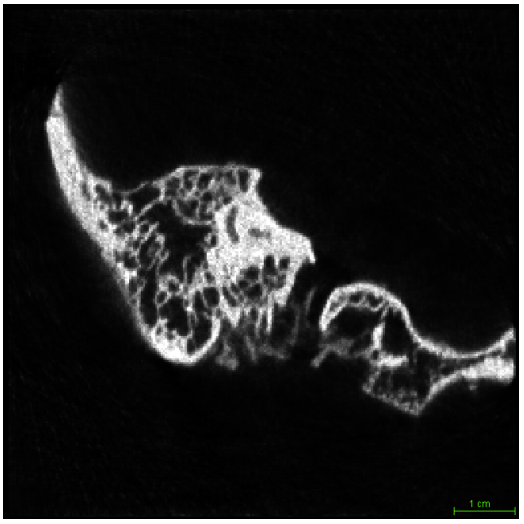
(b) FDK



(c) SART



(d) TV



(e) Proposed

Figure 3.55.: Reconstruction of the temporal bone with different reconstruction methods.

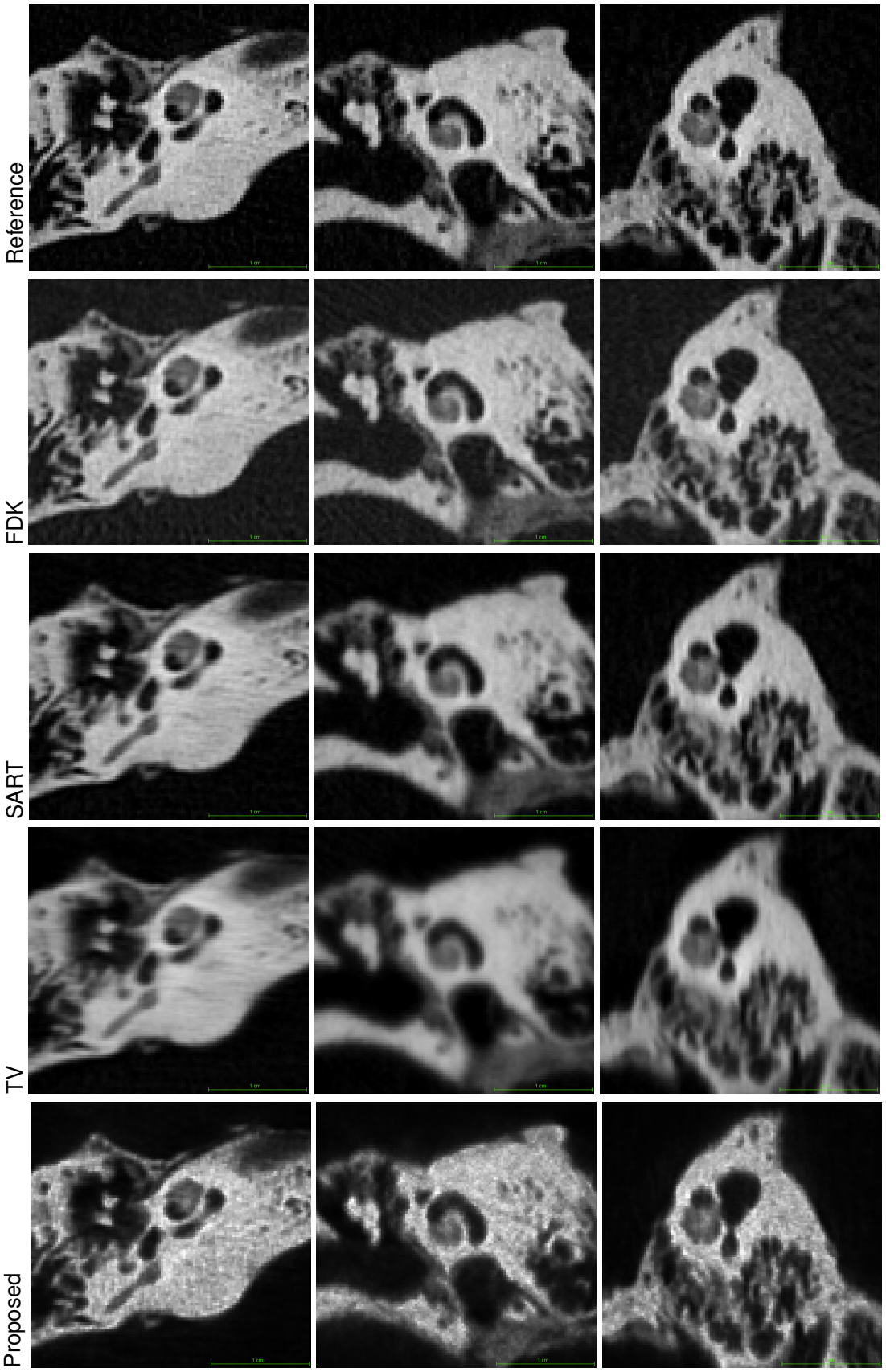
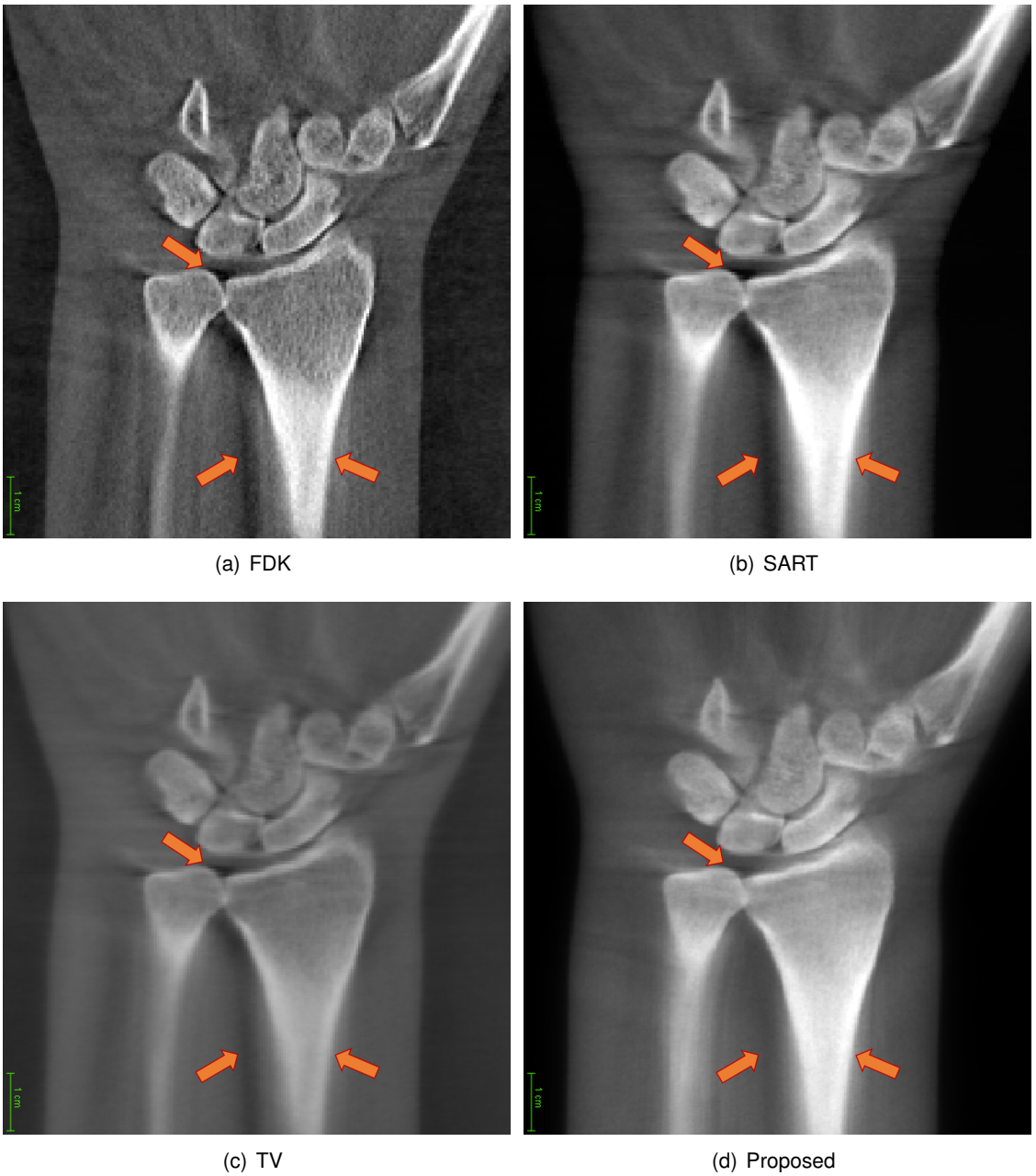


Figure 3.56.: Reconstruction of the temporal bone with different reconstruction methods. The left column contains the axial views, the middle column the coronal views, and the right column the sagittal ones.

Real CBCT data. The Figures 3.57 and 3.58 show coronal and axial images of the reconstructed hand phantom. The X-ray images were acquired with the C-arm CBCT system under investigation (see Section 3.2.1). It can be seen that the streak artifacts, especially the dark streaks (shadows), are substantially reduced using the proposed iVar-based reconstruction. Furthermore, the contrast to the soft tissue is enhanced and the artifacts in the background are significantly reduced.



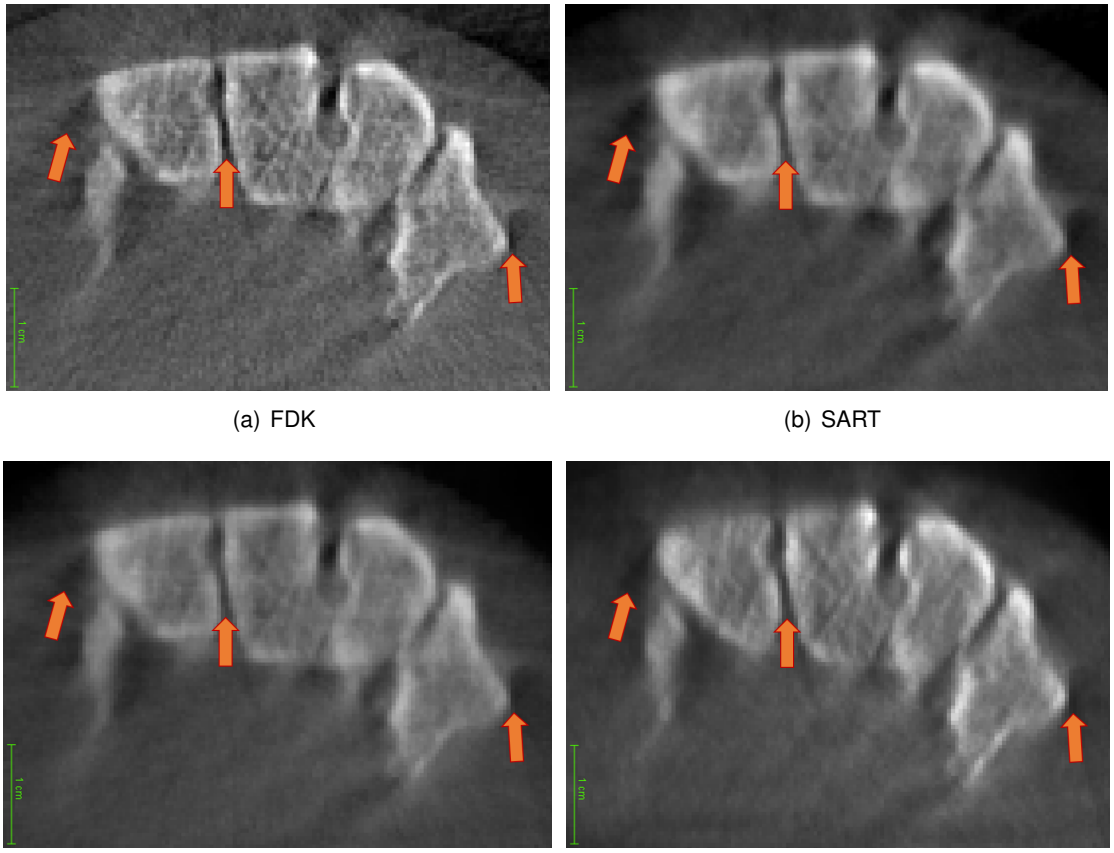
(a) FDK (b) SART

(c) TV (d) Proposed

Figure 3.57.: Reconstruction of a hand phantom scanned with C-arm CBCT system (coronal view).

Conclusion. It was shown that the reconstruction quality was substantially enhanced by applying the proposed CBCT reconstruction algorithm. However, it was not possible to completely remove the metal-related artifacts. For instance and as illustrated in Figure 3.53, dark as well as bright strikes are still

slightly visible between the metallic spheres. These artifacts were significantly attenuated compared to the other reconstruction methods, however not completely removed. In the next section, the causes of such remaining artifacts are explained and an extension of the proposed algorithm is proposed in order to completely eliminate also these CT artifacts.



(a) FDK (b) SART

(c) TV (d) Proposed

Figure 3.58.: Reconstruction of a hand phantom (wrist) scanned with C-arm CBCT system (axial view).

3.7. Metal artifact reduction

If the object under investigation contains metallic parts, some regions of the acquired X-ray images are useless for the CBCT reconstruction, since the rays attaining the corresponding detector cells are obscured by metals, i.e. recorded X-rays are highly attenuated. Usually, the correction terms calculated at such pixels are wrong, since the detector response is not anymore proportional to the total attenuation along the corresponding projection ray. After back-projecting such inaccurate terms, the air and soft tissue voxel which are crossed by the said projection ray would be significantly altered. If these inappropriate voxel updates are not correct by the next processed X-ray images, metal artifacts are created. This results in excessive metal artifacts as well as lost of details/structures in the reconstructed CT volume.

The algorithm proposed in Section 3.6 considers all pixels over all X-ray images, although pixels where metallic parts are projected are not appropriate to estimate the values of non-metallic voxels. In order to overcome this problem, two alternatives could be applied to reduce the metal artifacts:

- skip the (X-ray) pixels where air and soft tissue voxels are obscured by metal-voxels. For instance, use only the X-rays acquired within the green arc (see Figure 3.59) to reconstruct the air and soft tissue voxels. An initial volume estimate has to be reconstructed in order to classify the pixels corresponding to a given artifact voxel in appropriate (green) and inappropriate (red) pixels. After that the current CBCT volume should be updated while considering only the green pixels to reconstruct the artifact voxels.
- make the images acquired over the red arc in Figure 3.59 also appropriate for the reconstruction of the artifact voxels. This can be achieved by considering the metal-pixels (where metal parts are projected) as missing information and try to recover them, so that metal-free X-ray images are estimated from the acquired raw data.

The second alternative is straightforward and computationally inexpensive, but it requires at a given stage an accurate segmentation of the metal pixel in the X-ray images. Good segmentation results can be achieved if the segmentation is performed in the reconstruction space, i.e. based on the CT volume, since the metal implants can be better localized there, namely based on their density values which lie in a known range (>3000). However, this is not robust and accurate enough, since the metal artifacts voxels, especially the bright streaks, have values similar to metal voxels. For a robust and accurate segmentation we propose to employ the VMRs . Accordingly, the second alternative is followed in this thesis.

The previously proposed VMR-based reconstruction algorithm is extended to eliminate metal-artifacts by performing the following steps:

1. Reconstruct an initial CT volume V_{org} employing the originally acquired X-ray images P_{org} (see Section 3.6)
2. Detect “metal voxels” (see Section 3.7.1)
3. Project only metal voxels on all X-ray images and create mask for metal pixels
4. Generate metal-free X-ray images P_{cal}
5. Reconstruct a second CT volume V_{cal} based on the estimated metal-free X-ray images P_{cal}
6. Reconstruct the final CT volume V_{final} by blending densities from V_{org} and V_{cal} based on VMR-values for every voxel.

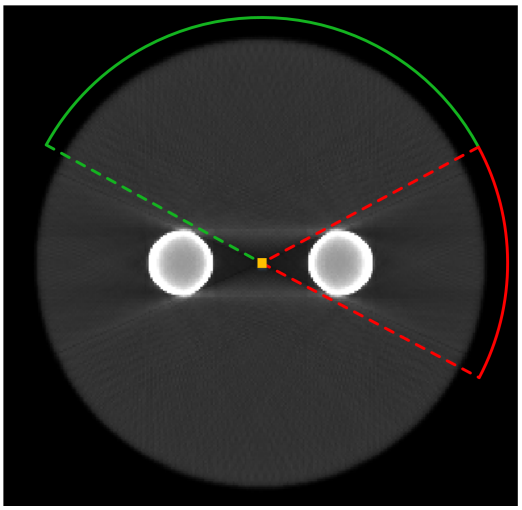


Figure 3.59.: Illustration showing a voxel (yellow box) which is affected by metal-artifacts. In some gantry angles (red arc) this voxel is obscured by the metallic spheres on the X-ray images, unlike the rest of projections (green arc).

In the following, the second, third, fourth and sixth processing steps are described with more details. The 3D reconstruction performed in the first and fifth steps is done using the proposed VMR-based algorithm presented in Section 3.6.

3.7.1. Segmentation of metal pixels

Performing the segmentation of metal pixels directly on each X-ray image is a challenging task even for manual segmentation. For illustration, an example is shown in Figure 3.60. The X-ray image is pseudo-colored to emphasize the contrast problem preventing the successful segmentation of the metal objects with a simple thresholding. Effectively, it is difficult to segment the dental braces, especially in the regions with low contrast.

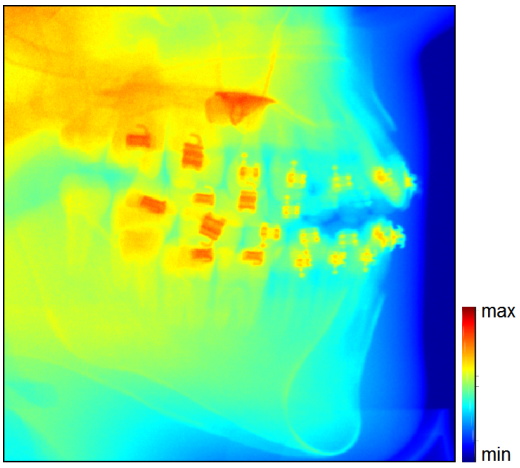


Figure 3.60.: An example illustrating the difficulty to segment metal objects in 2D X-ray images. Pseudo colors are employed to emphasize the contrast problem.

The simplest segmentation method is to extract as metal voxels whose intensities exceed some threshold [Glo81, LMK*12]. However, this approach can be employed only in case of artifacts-free reconstructions, since artifact regions can also have high intensities and would therefore be classified as metal voxels. In such a case an over-segmentation will result in degradation in the reconstruction quality and induce new CT artifacts. More sophisticated techniques have been proposed to enhance the segmentation results, e.g. by incorporating local region information [KCWM12]. These techniques do not necessary perform well, since the artifacts are not always spatially close to the metal voxels, but also striped in the whole volume. The last category of segmentation approaches for the mentioned purpose defines some models to assist the segmentation [BS06]. These algorithms are usually more accurate, but they require prior knowledge about the scanned object, and can therefore not be used to segment unlearned (unknown) objects.

3.7.2. Detection of metal voxels

This thesis contributes a new algorithm for accurate segmentation of “metal pixels” without requiring prior knowledge about the object to be reconstructed and/or the metal parts present in the scanned field. A way to overcome this problem is to preform the segmentation of metal voxels in an initially reconstructed CT volume. If the metal detection is robust and reliable, find the projection areas (pixels) corresponding to the

segmented metal regions. Typically, metal voxels are characterized by their high intensity, but also voxels belonging to bright regions of metal artifacts have high intensities. In order to distinguish between these two voxel types, the variance-to-mean ratio can be employed (see Figure 3.38(b)): artifact voxels show high VMR values, unlike metal voxels which are characterized by their low VMRs. These two characteristics (intensity and VMR) are exploited to accurately extract metal voxels.

Let M_{metal} be a binary segmentation mask defined as

$$M_{metal}(i) = \begin{cases} 1 & \text{if } V(i) \geq \tau_{metal} \wedge VMR(i) \leq 1.0 \\ 0 & \text{else} \end{cases} \tag{3.29}$$

where τ_{metal} is a threshold defining the minimum volume density of "metal voxels" in the reconstructed image ($\geq$3000).

After segmenting the metal voxels and based on the corresponding geometrical projection parameters, the center points of the metal voxels should be projected on the 2D projection plane (detector plane). A projection pixel is defined as a "metal pixel" if at least one "metal voxel" (mask value =1) is projected on this pixel position. However, in order to avoid gaps in the projection regions (on the X-ray image) we iterate over all pixel and check for each pixel if the corresponding projection ray intersect with at least one metal voxel.

3.7.3. Metal-free X-ray images

After labeling the metal objects in all X-ray images, all pixels which have a mask value equal to 1 are assumed to be metal pixels and have to be replaced in order to create metal-free X-ray images. The metal pixels are considered as gaps of missing information which have to be filled in a meaningful way. A bilinear interpolation of neighbors over masked pixels may deliver satisfactory results only if the orientation of structures on the X-ray images is horizontal or vertical. For this purpose an inpainting technique has been applied. Inpainting methods based on PDE [BSCB00] or statistical techniques [BO07, OB09] may perform well, but they are very time consuming and are therefore unpractical for clinical applications. A comprehensive approach is the directional interpolation [AFP91]. This algorithm aims to keep definite geometric structures, such as edges, and to interpolate in the low frequency direction, i.e. along edges. This approach was extended to fill the gaps as proposed in Section 2.4.5. An example of estimated metal-free X-ray image is shown in Figure 3.61. It can be seen that details are continuously reconstructed in the replaced pixels.

3.7.4. Fusing CTs

At this level two CT volumes are reconstructed. The first volume V_{org} is calculated based on the original X-ray images. It contains therefore the metallic parts and obviously the related metal artifacts. The second volume V_{cal} is calculated based on the metal-free X-ray images. It contains no metal voxels and consequently no metal artifacts. Generating the final CT volume by strictly copying the corresponding voxels from either V_{org} or V_{cal} may result in a synthetic appearance of the reconstructed volume. Here it is important to take all objects and details into account while discarding artifacts. In order to overcome this problem, a weighted combination of V_{org} and V_{cal} has been applied. For this purpose, VMR values, which represent

(a) Acquired X-ray image (b) Estimated metal-free image

Figure 3.61.: An example illustrating the ability to accurately segment metal pixels in an acquired X-ray image (a) and replace the segmented pixel in order to create an appropriate metal-free projection image(b).

the probability of each voxel to be a part of an artifact region, are considered. Voxels with low VMRs, i.e. metal and artifact-free voxels, are mainly influenced by V_{org}. In contrast, artifact voxels with high VMRs are influenced by V_{cal}. The fusion function is formulated as

$$V(i) = \eta \cdot V_{org}(i) + (1 - \eta) \cdot V_{cal}(i) \tag{3.30}$$

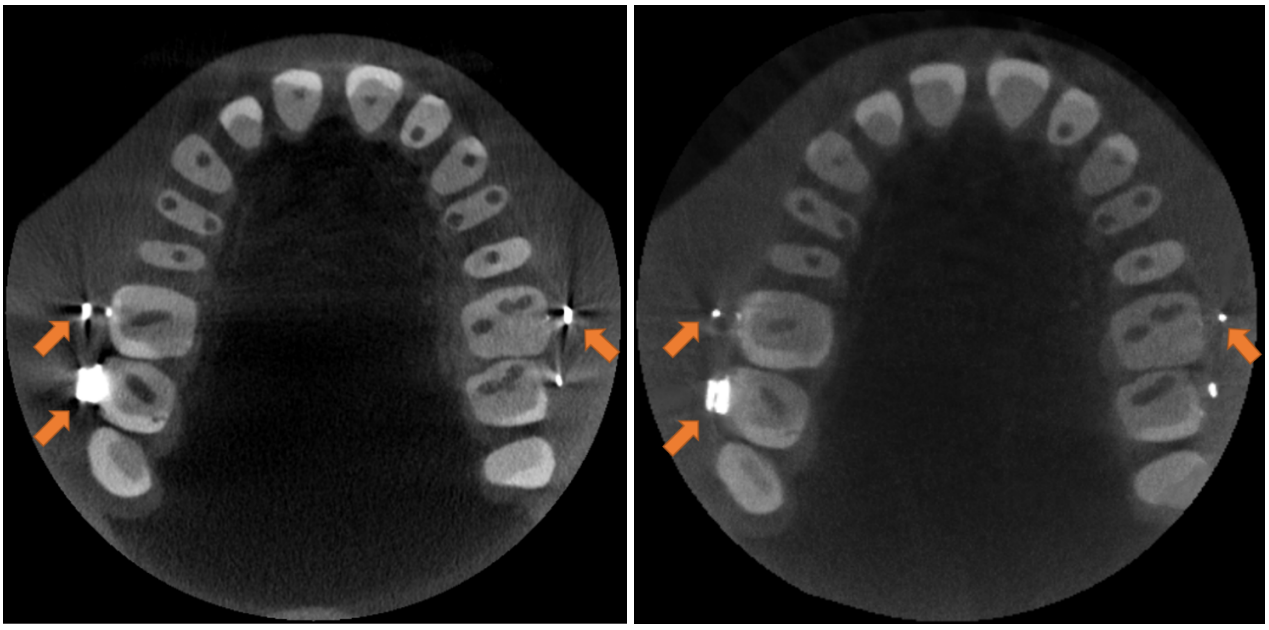
where η serves for normalization: the VMR of each voxel $VMR(i)$ is scaled relative to the maximal VMR (VMR_{max}))within the whole CT volume

$$\eta = \frac{VMR(i)}{VMR_{max}} \tag{3.31}$$

3.7.5. Evaluation of metal artifact reduction

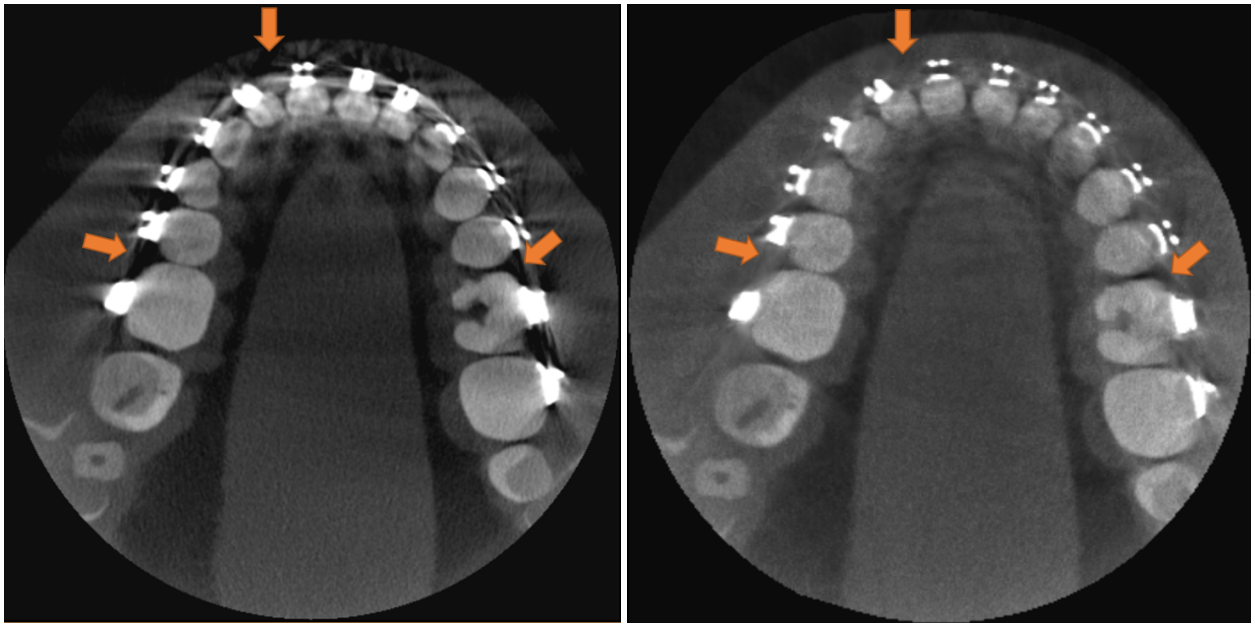
A frequent problem in the head and neck CTs is the presence of artifacts caused by dental implants. Especially when traces of high-contrast structures are present in the scanned regions, severe CT artifacts between metal implants and teeth are induced. This is a challenging case for testing the ability of the proposed algorithm to deal with artifacts caused by dental hardware. The X-ray images used in these tests are acquired with a Planmeca 3D CBCT imaging device (ProMax 3D). As shown in Figure 3.62 and Figure 3.63, the metal-related artifacts are completely suppressed. Ghosting artifacts in the oral cavity and in the Gingiva are significantly reduced.

Plates and screws are usually used in orthopedic surgery for fixation of bone fractures. A representative scan of a hand with a dorsal locking plate and screws is shown in Figures 3.64 and 3.65. The streak and shadow artifacts caused by the metal parts are significantly reduced using the proposed method.



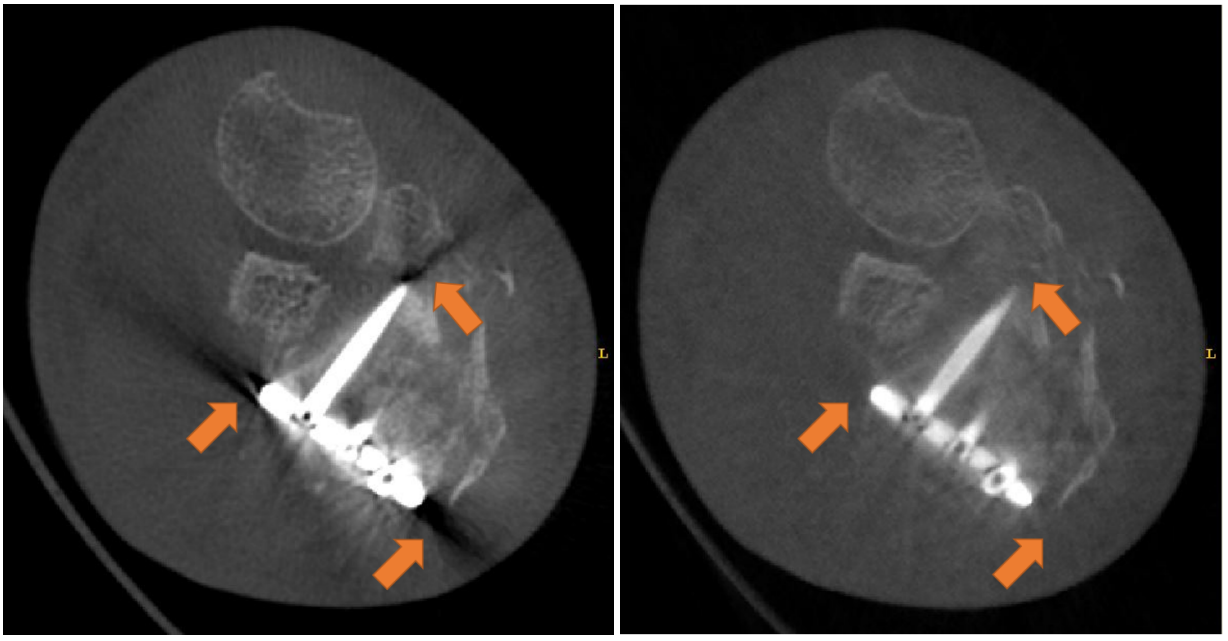
(a) FDK (b) Proposed

Figure 3.62.: Artifacts originating from dental brackets (indicated by arrows) are completely suppressed by 3D reconstruction algorithm.



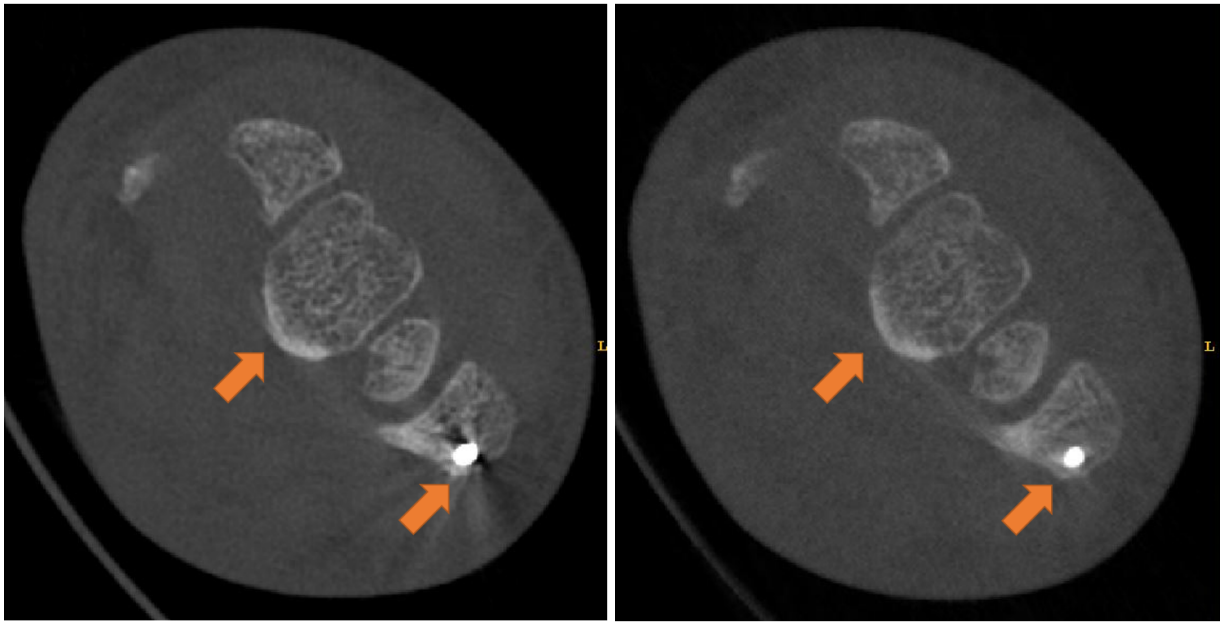
(a) FDK (b) Proposed

Figure 3.63.: Artifacts originating from dental brackets (indicated by arrows) are completely suppressed by 3D reconstruction algorithm.



(a) FDK (b) Proposed

Figure 3.64.: Artifacts originating from metal parts (indicated by arrows) are completely suppressed by the proposed 3D reconstruction algorithm.



(a) FDK (b) Proposed

Figure 3.65.: Artifacts originating from metal parts (indicated by arrows) are completely suppressed by the proposed 3D reconstruction algorithm.

Computational complexity. The tests were performed on an Intel Core i5-4670 CPU with 3,40 GHz CPU and 16 GBytes RAM. The used graphics card was a GeForce 7600 GTX. The size of the CT images and the X-ray images was $512 \times 512 \times 512$ and 1024×1024, respectively. The processing times reported below are computed for 1 iteration as well as 3 iterations (which is sufficient to achieve good reconstruction quality).

Operation	Times (in seconds)
pre-filtering and computing VMRs	6.24
forward projection	9.18
inpainting	2.8
computing correction terms	0.63
back-projection	4.33
total (1 iteration)	23.18
total (3 iterations)	57.06

Table 3.5.: The execution times needed for the CBCT reconstruction.

3.8. Summary and conclusions

In this chapter several new algorithms have been proposed to deal with two main CT artifacts sources, namely the mechanical inaccuracies and the presence of high-density (metal) objects in the scanned regions, which results in typical metal-related artifacts.

An iterative on-the-fly calibration for a mobile C-arm CBCT system has been presented. The importance of each geometric projection parameter and its influence on the reconstruction quality have been studied in order to eliminate the parameters causing only minor degradation in the image quality. This served to simplify the calibration process and make it more robust as well as less computationally complex. Effectively, the calibration could be reduced to an in-plane registration without significant degradation in the image quality. This simplification is valid at least for our specific C-arm device, other scanner geometries might require to consider additional geometric projection parameters. Experiments based on simulated and real cone-beam data have demonstrated the ability of the proposed calibration method to achieve sub-pixel accuracy.

A novel non-linear back-projection kernel for iterative CBCT reconstruction has been proposed to reduce CT artifacts and enhance the soft tissue contrast in case of moderate metal-related artifacts. The introduced variance-to-mean ratio was employed to detect the presence of metal artifacts. If this ratio exceeds 1.0, heavy metal artifacts are usually intended in the reconstructed volume. This criterion can serve to decide about the use of the variance-based reconstruction in its extended version (with metal-artifact reduction) or not. An extension of this algorithm has been also presented to deal with excessive metal artifacts and recover details lost in the corrupted regions. Evaluation performed with simulated and real data have shown that CT artifacts could be substantially reduced by applying the proposed reconstruction approach. Metal artifacts could even be completely eliminated using the extended version of the proposed technique. Despite the computational complexity of the presented algorithm, it is still fast enough to be suitable for clinical applications (<1min).

4. C-Arm image-guided navigation

4.1. Motivation

Any significant deviation from the planned trajectory during drill process may result in damaging critical anatomical structures. It is therefore mandatory to intra-operatively control the drilling by measuring the current drill pose (position and orientation) of the drill bit and then compare it with the planned drilling path. In case of any abnormalities, the drill process has to be stopped and appropriate decision and/or correction has to be made by the surgeon. In this context the image-guided navigation can be used to control the drilling.

Image guided intervention (IGI) covers a wide range of surgical and therapeutic procedures [CP10], such as image-guided minimally invasive surgery. This later provides information about the current position of the surgical instrument relative to the planned trajectory [PDBF12]. All IGI approaches rely on registration and/or tracking techniques to align the current surgical instrument position to the patient and the planning data.

In recent years C-arm CBCT systems have been used for image-guided interventions [Sie11]. It has been frequently recommended to employ intra-operative CBCT volumes for navigation due to their high accuracy compared to fluoroscopy guidance since it is easier to identify the surgical instrument in CBCT volumes as in (2D) X-ray images [CKK*07]. However, repetitive acquisition of fully 3D CBCTs results in excessive radiation exposure for the patient. An alternative to significantly reduce the radiation dose is the use of only two X-ray images, e.g. lateral and oblique views, instead of CBCTs. At least two X-ray images are needed in order to recover the missing depth information, because a single X-ray image lacks 3D spatial information. For this task photogrammetry methods can be applied to compute the current pose of the surgical instrument: The spatial position of a given target point, e.g. the tip of the drill bit, is the intersection of the X-ray projection lines passing through the projection location of the said point on each of the X-ray images as illustrated in Figure 4.1. After calculating the spatial pose of the drill bit, its position and orientation is compared to the planned drilling path.

Using the C-arm CBCT system exclusively in the fluoroscopic mode for navigation is associated with several challenges, especially the accurate estimation of the C-arm pose relative to the patient. The pose estimation has to be fast, accurate and robust to be appropriate for clinical applications. A favorable alternative which could satisfy all of these requirements is the marker-based registration which employs fiducial markers to estimate the C-arm pose [PDBF12]. It is well known that the target registration accuracy of such an approach depends on the fiducial localization error [FW01]. Thus, it is very import to localize the fiducial markers in CT volumes as well as in X-ray images with high accuracy.

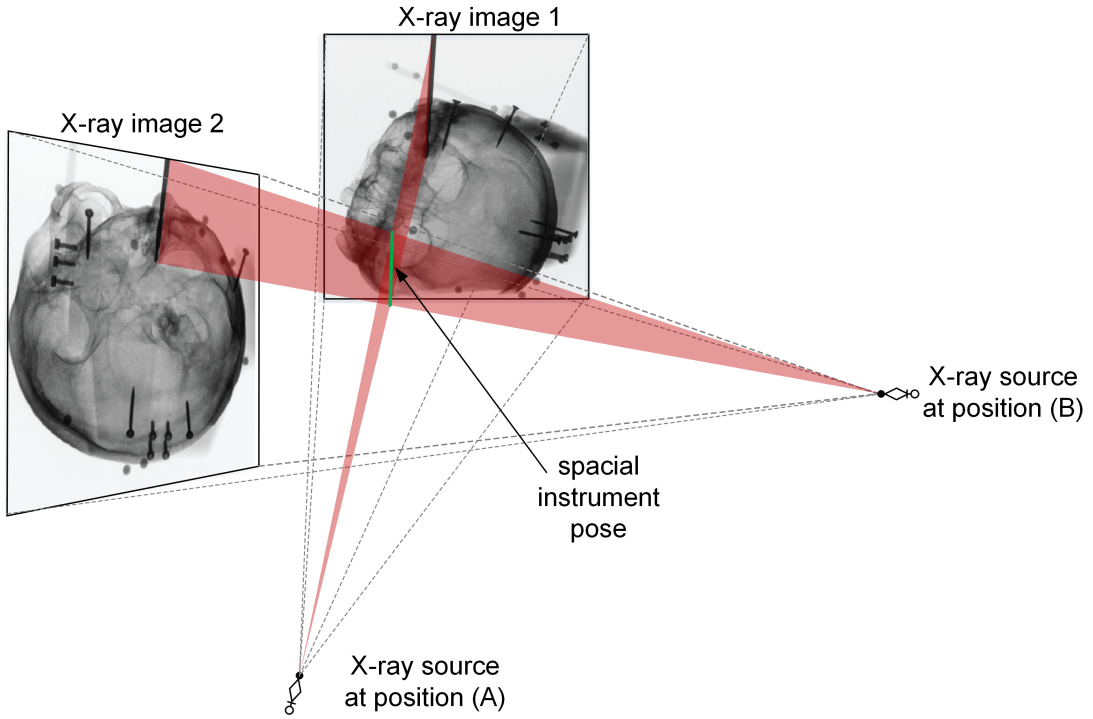


Figure 4.1.: Illustration of the disposition of two X-ray images aiming at estimating the 3D position of the surgical instrument (here the drill bit).

Navigation concept

In a proxi-operative stage a base plate (shown in Figure 4.3(a)) is fixed to the patient skull. This plate contains fiducial markers which are employed to align the X-ray images to the patient's anatomy. The used base plate is of small size and can therefore be fixed to the patient's head without limiting his mobility. In such a way it is possible to perform the pre-operative CT in a separate station, and than move the patient later to the operation room. Therefore, it is not a prerequisite anymore to place the pre-operative and the intra-operative imaging systems in the same room, which would allow an optimal utilization of available resources. For instance, a pre-operative CT can be acquired with a conventional spiral CT, then the patient would be moved to an operation room equipped with a C-arm CBCT system. Between the two stations the base plate should not be removed or shifted from its position relative to the skull since this would hamper the registration step.

The pre-operative CT volume is provided as starting point for the image guided navigation which consists of the following components:

1. Planning: Critical structures are segmented in the pre-operative CT, then a simulation is performed to automatically determine feasible drill paths and select optimal combination of three feasible drilling paths. Methods allowing these tasks have been developed by Meike Becker as a part of her dissertation [Bec14].
2. C-arm pose estimation: Establish spatial transformation between the intra-operatively acquired X-ray images and the pre-operative CT with high accuracy. This step is the focus of this chapter and is described in Section 4.2.
3. Instrument localization: Determine the position and orientation of the surgical instrument relative to the patient and the planned trajectories. An image-based tracking approach aiming to localize the

drilling instrument in X-ray images has been developed under my supervision by Beatrice Friess within her Master's theses [Fri14]. A single X-ray image lack 3D spatial information, thus it is mandatory to recover the missing depth information by acquiring at least two X-ray images in order to determine the spatial location of the surgical instrument (drill bit). After localizing the drilling instrument (centerlines) in both intra-operative X-ray images, the pose of the drill bit relative to the patient and planning data is calculated by employing the geometric projection parameters estimated in step 2 (see Section 4.3).

4. Navigation: Compute the deviation of the drill instrument from the planned trajectory and visualize this in order to asses the drill process (Section 4.4).

The interaction of the these components is illustrated in Figure 4.2.

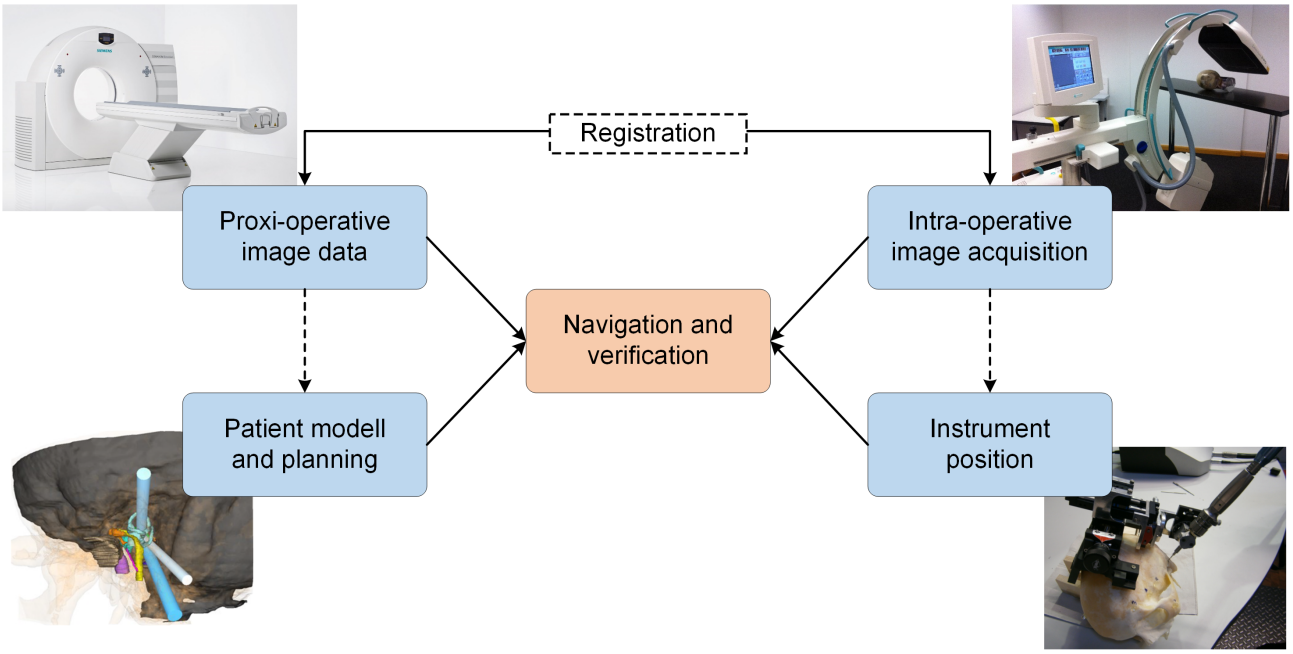


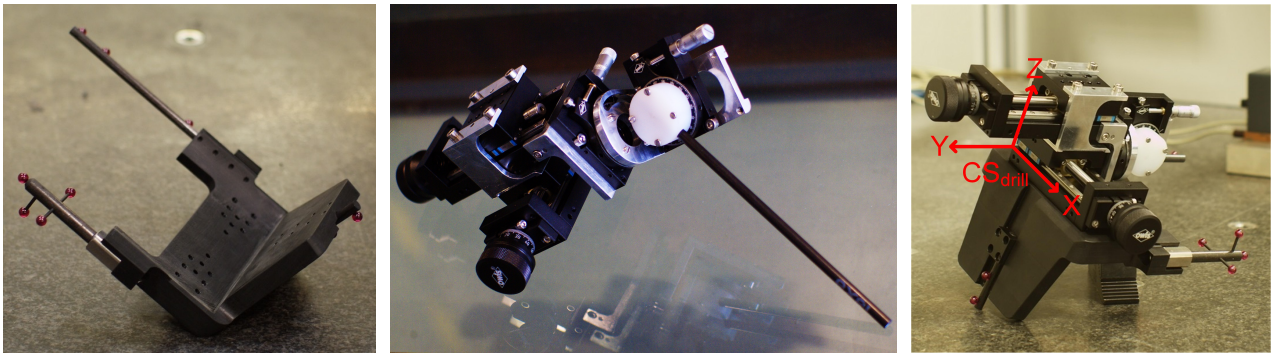
Figure 4.2.: Illustration of the interaction between the different modules of the image guided intervention. The Figure of the spiral CT scanner is from [Sie]

The reader is referred to [Bec14] for all details related to the first step of the navigation pipeline. In the rest of this section some background information about the employed base plate, the choice of the marker material, and the used coordinate systems are presented.

Base plate

A prototype has been developed within the MUKNO project for positioning the drill instrument [NPS13]. It is designed to be fixed to the skull with surgical screws and support the drilling system as depicted in Figure 4.3(c). Unlike most of positioning devices which have been proposed for skull-base surgery, such as the stereotactic frame [KNY*11], our base plate has several advantages:

- it is of small size
- it does not have to be individual for each patient
- it is made of Polyoxymethylene (POM) in order to avoid CT artifacts since such a material has low attenuation coefficient



(a) Base plate (b) Positioning system (c) Positioning system attached to the base plate

Figure 4.3.: Base plate and positioning system used to drill the planned drilling paths. Fiducial markers are fixed to the base plate (red spheres) for registration purposes.

Mechanical positioning system

The drill positioning system shown in Figure 4.3(b) is used to align the drill bit according to the planned trajectory. It consists of two linear and two rotatory axes. The linear axes have a translational course of 45 mm, a pitch angle of max. 250 µm, and a yaw angle of max. 300 µm with vertical and lateral deviations of max. 7 µm and 10 µm, respectively. The angular adjustment is in the range of 360°. The drill guide can be manually adjusted within a 35×35 mm^2 area. The device is equipped with brakes which can be locked in order to prevent axis movements after setting the desired axis values.

Fiducial markers

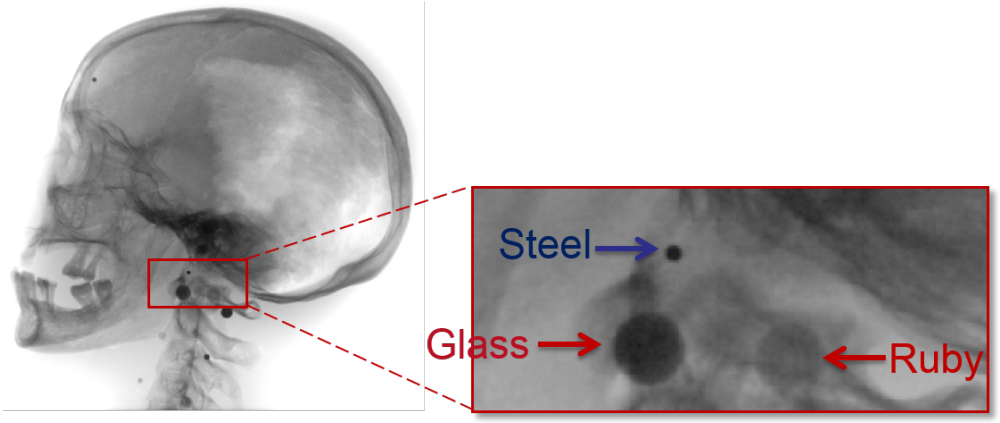
Reference structures (fiducial markers) are fixed to the base plate. They can be detected in CT volumes as well as in X-ray images and can therefore be employed to :

- register the planning data to the local coordinate system of the positioning device
- estimate the current C-arm pose during the drill process
- align the drilled tunnels relative to the planned drilling paths

The marker configuration, i.e. their distribution and placement relative to the region of interest, has to be chosen in a way to maximize the registration accuracy. In addition, the choice of the marker material and size is very important since it influences the registration accuracy as well as the CT reconstruction quality. For instance, metal markers can be easily detected in X-ray images and are therefore advantageous for the registration, but they would obscure the critical anatomical structures and cause CT metal artifacts. In contrast, plastic markers would not cause any CT metal artifacts, but they are heavily detectable and not accurately localizable in the X-ray images, which would increase the target registration error. The marker size is also important since it influences the resulting attenuation value and consequently the detection accuracy. For example, large markers made of glass appear darker in the X-ray images compared to small glass fiducials.

In order to determine appropriate proprieties (material and size) for skull-base surgery, several markers made of different materials and with different radii were place on a cadaver skull phantom and scanned at different energy levels with CT and X-ray images. An example of acquired X-ray images is shown in

(a) Spherical fiducial makers which are made of different materials (metal, steal, glass, PA, POM, and PTFE) and have different radii (1mm to 10mm)



(b) X-ray image of different spherical markers attached to a cadaver skull phantom

Figure 4.4.: Determination of appropriate marker proprieties (material and size) for skull-base surgery.

Figure 4.4. From our experience glass markers with radii in the range of 4 mm to 6 mm are appropriate for the investigated problem since they do not obscure the rest of anatomical structures in X-ray images while keeping a relatively good contrast to the surrounding region. Alternatively, fiducials made of synthetic ruby can be used, but they have lower contrast compared to glass markers and are therefore more difficult to be localized in X-ray images..

Coordinate systems

The coordinate systems used in the navigation are illustrated in Figure 4.5:

- CS_{CT}: the local coordinate system of the CT volume
- CS_{world}: the world coordinate system defined by the fiducials on the base plate which is fixed to the patient's skull
- CS_{drill}: the local coordinate system of the positioning (drilling) system
- CS_{XC}: the local coordinate system of the C-arm CBCT system.

Accordingly, these transformations are defined:

- $\mathbf{T}_1$: transforms the planned paths defined in CS_{CT} to the world coordinate system CS_{world}. It is computed based on the coordinates of the marker centers in the CT volume determined as described in Section 4.2.3.
- $\mathbf{T}_2$: transforms the planned paths from CS_{world} to the local coordinate system of the positioning system CS_{drill}. In order to accurately determine $\mathbf{T}_2$, the fiducials markers as well as characteristic points on the guide rails are scanned with a coordinate measuring machine CMM (ZEISS UPMC

850 CARAT and Prismo Vast, Fa. Zeiss, Germany), as shown in Figure 4.6. The transformation $\mathbf{T}_2$ is then easily computed based on the point correspondence since all points are measured in the (same) local coordinate system of the CMM.

- $\mathbf{H}_1$ and $\mathbf{H}_2$: transform the position of the tracked instrument relative to the C-arm, i.e. from P_1 or P_2, to the CT volume (CS_{CT}). They are calculated by employing the C-arm pose estimation described in the next section.

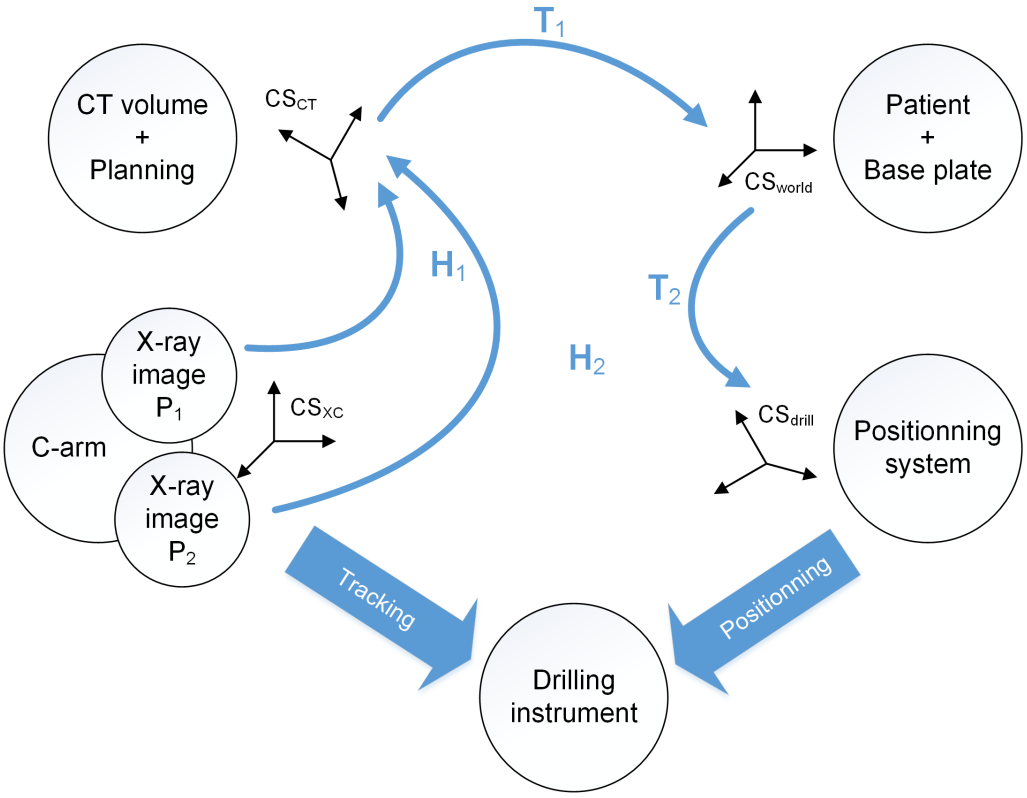


Figure 4.5.: Illustration of the used coordinate systems.

4.2. C-arm pose estimation

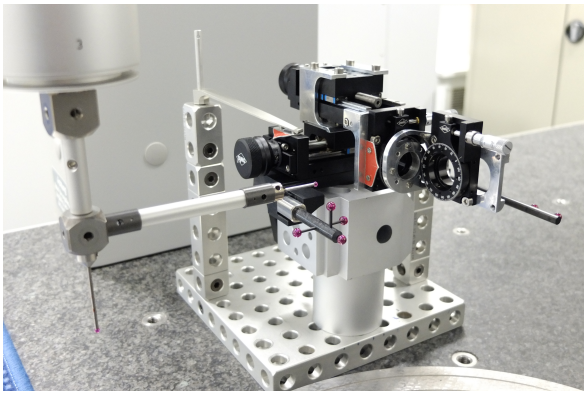
The pre-operative data set is a CT volume. The intra-operative data are either X-ray images or CBCT volumes. According to the dimensionality of the intra-operative images, it is to be distinguished between 2D-3D and 3D-3D registration. Commonly, intra-operative CBCTs are (rigidly) aligned to the pre-operative CT volume in form of paired-point registration by employing anatomic points or fiducial markers.

In this thesis the provided data sets are a pre-interventional CT volume and intra-operative X-ray images. Therefore, the focus is set in the following on 2D-3D registration.

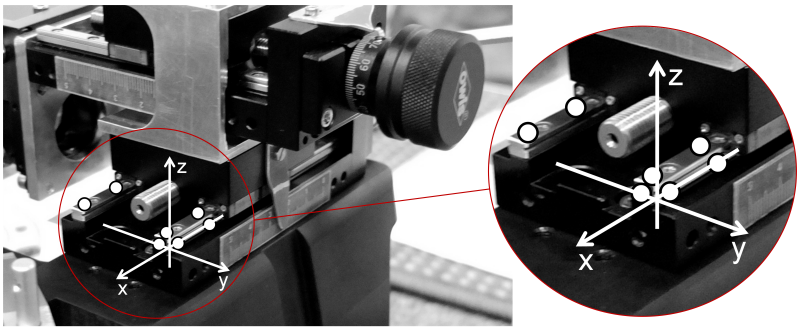
4.2.1. Related work

For a successful image-guided navigation it is mandatory to accurately determine the pose of the C-arm relative to the patient anatomy and the associated surgical plan. This step is performed in form of 2D-3D

(a) CMM



(b) Measurement of the marker positions



(c) Local Cartesian coordinate system of the drilling system CS_{drill} defined based on characteristic points on the guide rails

Figure 4.6.: Coordinate metrology machine (CMM) used to accurately determine the position of the fiducial markers relative to the positioning device.

image registration which matches the intra-operative X-ray images to the pre-operative CT volume. The 2D-3D image registration methods can be classified as:

1. calibration-based methods: the imaging device is pre-calibrated so that its position relative to the patient can be derived from the calibration protocol.
2. intrinsic approaches which rely only on the patient's anatomy, i.e. no additional foreign fiducials are used. These methods can be further divided in three groups: feature-based, gradient-based, and intensity-based registration.
3. extrinsic approaches which employ fiducial markers implemented to the patient in order to achieve the spatial alignment. The fiducials are designed to be detectable in CTs as well as in X-ray images.

The last two categories are of interest for this thesis and are therefore presented in more detail in this section, whereas the focus is set on marker-based (extrinsic) registration for C-arm systems. For more details on the several registration techniques the reader is referred to [Ste14] and [PDBF12].

1) Calibration-based registration. The imaging device has to be pre-calibrated, so that its position relative to the patient table can be derived during the image acquisition. In order to establish the position of the operation table relative to the C-arm, the different components of the system, i.e. the C-arm and the X-ray table, have to be tracked, e.g. using an optical tracking system with infrared emitting diodes. Assuming that the relative pose of the patient relative to the operation table remains unchanged, the current patient pose can be calculated. For instance, this technique was applied to align X-ray images to MR

images [MRA*06]. This approach was also proposed for advanced workplace for image-guided surgery (AWIGS) [BBB*07]. This later is equipped with an operating table which can transfer the patient into the CT scanner. The reported calibration accuracy is 0.7 mm with a bias of 2.1 mm. Limitations imposed by this technique are the immobilization the patient during the intervention and the use of additional devices, such as the tracking systems, which make this approach only conditionally appropriate for clinical application. Moreover, the achieved accuracy is not sufficient for the investigated problem. Therefore, this registration category is not followed further in this thesis.

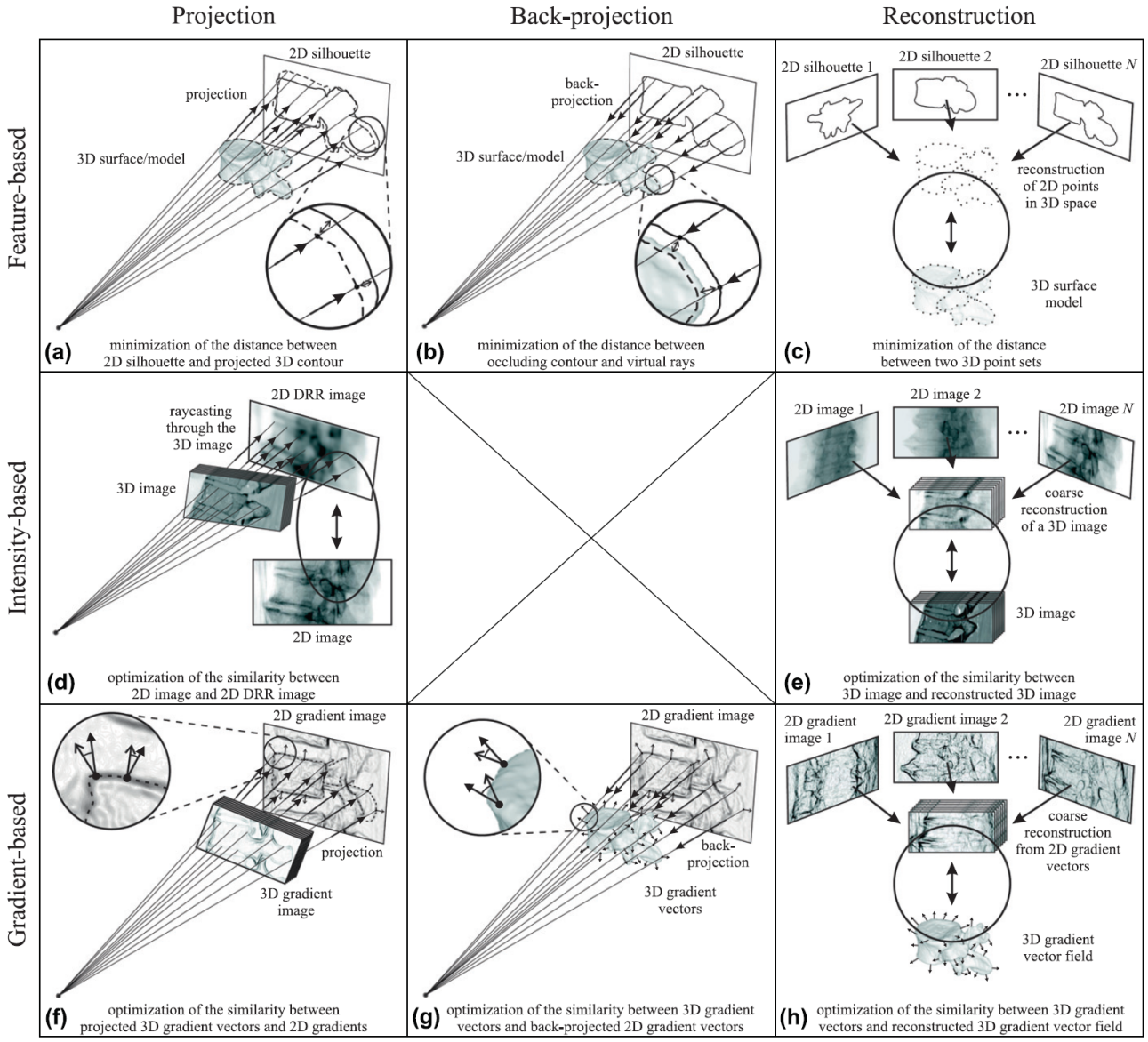


Figure 4.7.: Geometrical setup of intrinsic 2D-3D registration methods according to the nature of the registration basis and the strategy to achieve spatial correspondence [PDBF12].

2) Intrinsic 2D-3D registration. Here only anatomical structures are employed to perform the registration [PDBF12]. As depicted in Figure 4.7, this registration category is further classified as:

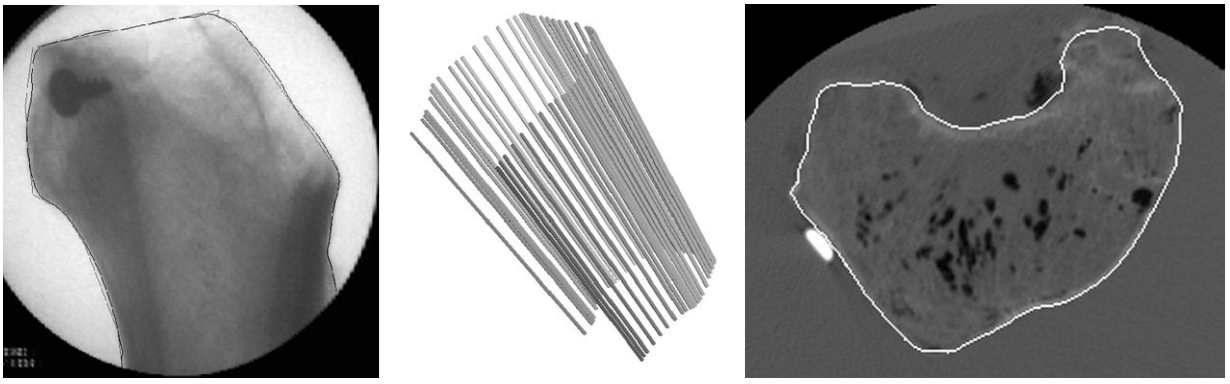
- feature-based registration

- intensity-based registration
- gradient-based registration

Feature-based registration aims to extract anatomical features in the pre-operative CT volume as well as in the X-ray images. After that, it tries to determine the transformation that minimizes the distance between the corresponding 2D and 3D features. Image segmentation is typically employed to extract geometrical features, however this would result in large registration error if the segmentation is not accurate. Gueziec and Kazanzides [GKWT98] proposed to extract bone surfaces and contours respectively from the CT volume and the X-ray images, then register the surfaces to the contours as illustrated in Figure 4.8. The extraction of anatomical structures is performed semi-automatically: the user have to select some points in the vicinity of the structures of interest as initialization for the adaptive active contour (snakes) segmentation. The authors have tested several methods for the registration task and recommended to use the apparent contours of the surfaces due to its advantageous computational complexity and robustness. The pose estimation is done by iterating on the following steps till convergence:

1. compute the apparent contours of the segmented surfaces
2. compute the closest point from anatomical surfaces to X-ray contours
3. update the geometrical projection parameters

The main drawback of this method is that inaccurate segmentation would result in additional registration errors. The authors reported that the the maximum registration error was between 1.2 mm and 3.7 mm. Such uncertainties are not acceptable in case of minimal invasive interventions.



(a) (b) (c)

Figure 4.8.: Illustration of the method proposed by Gueziec and Kazanzides [GKWT98]: register the surfaces extracted from the CT volume (a) to the contours from the X-ray image (c). The obtained X-ray paths are shown in (b).

Intensity-based registration tries to estimate the transformation by optimizing the similarity measure between the originally acquired X-ray images and the digitally reconstructed radiographs (DRRs) created based on the pre-operative CT volume. In this context a similarity measure is defined as a cost function for an optimizer which alters the actual projection parameters in order to achieve a better similarity value, i.e. a better alignment. The most used similarity measures are:

- (normalized) mutual information [GHv*01, PEH*07, LXQ13]
- (normalized) cross correlation [LBKH00, BSF*09]
- sum of square differences [JSRF06]
- entropy of difference image [PWL*98]

Further metrics have been proposed for local intensity changes, such as pattern intensity. The advantage of the intensity based registration is that it does not require any segmentation. Moreover, it exploits all the information in the images, therefore it is expected to be more accurate compared to the feature-based approaches. However, it has small capturing range and requires therefore initialization close to the searched pose in order to converge against the correct solution. Another drawback of this method is the high computational cost.

Gradient-based registration is performed by finding the best match between the 3D CT gradient vector field and the gradient projections on the X-ray images. First methods proposed in this registration category were introduced by Tomazevic et al. [TLSP03] and Livyatan et al. [LYJ03]. These methods are based on the observation that X-rays which pass through edges in the X-ray images are tangent to bone surfaces in the CT. The registration problem is solved by maximizing the sum of 3D gradient magnitudes which are incident on the bone edges in the X-ray image. An example illustrating this approach is shown in Figure 4.9. The reported target registration accuracy for this method is 0.5–1.7 mm, which is not accurate enough for lateral base surgery.

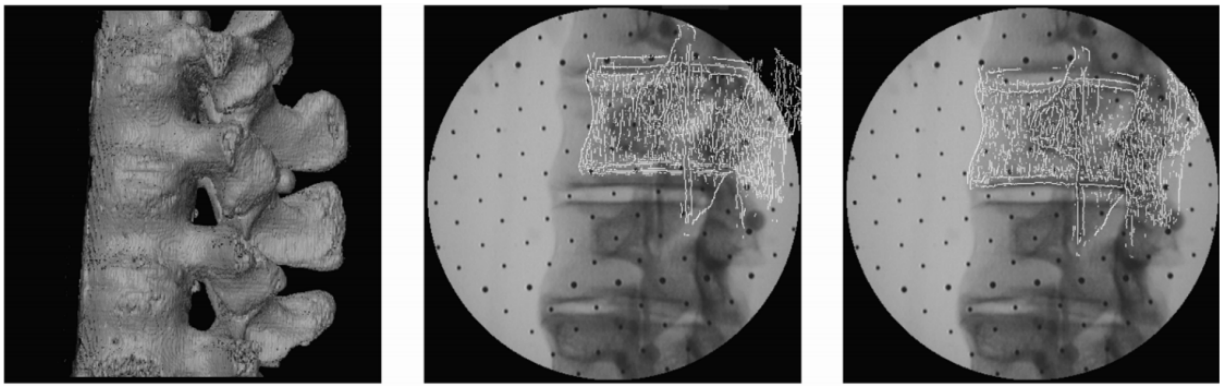


Figure 4.9.: Illustration of registration of dry vertebra with an X-ray image. The image on the left shows the CT model, the image in the middle shows the initial contours, and the image on the right shows the final pose superimposed on the X-ray image (white lines) [LYJ03].

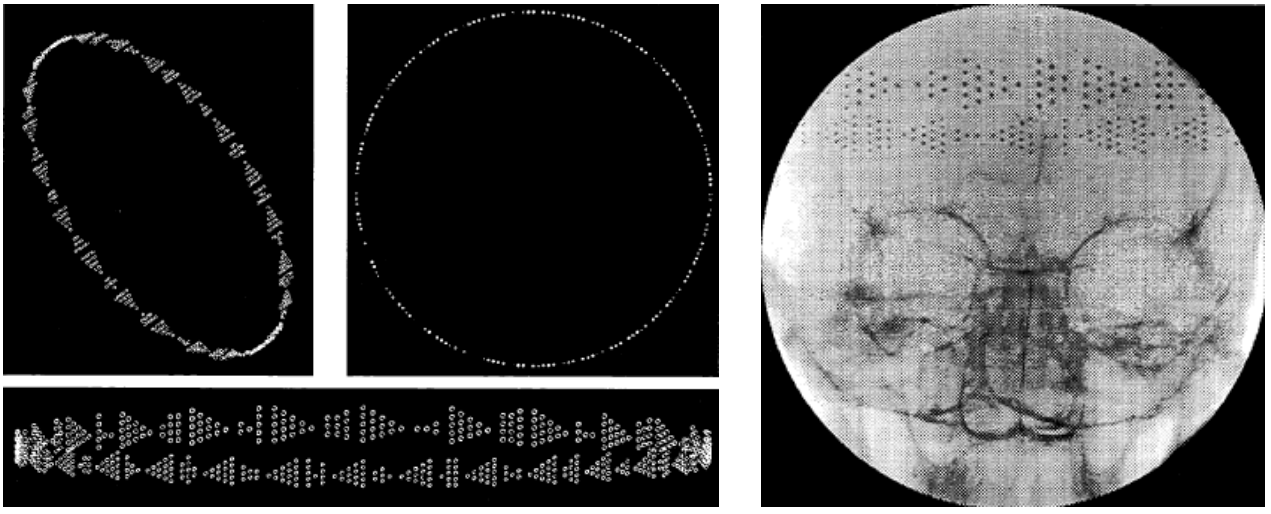
In summary, feature-based methods require to extract anatomical features and may therefore suffer from segmentation errors. This drawback is overcome by intensity-based methods since they do no require any segmentation. Furthermore, they are generally more accurate compared to other registration methods. However, they are susceptible to local optima and require therefore an initialization which is close to the optimal pose. This later problem is also noted in case of gradient-based techniques.

3) Extrinsic 2D-3D registration. Artificial objects or fiducials which are designed to be detectable in CTs as well as in X-ray images, like gold seeds, are fixed to the bone or soft tissue, and then used to perform the registration. This marker-based registration is fast and does not require complex optimization algorithms [PDBF12]. The most critical step here is the accurate localization of fiducials. It is to be noted that the markers have to be placed before acquiring the pre-operative CT, and must remain in their positions until finishing the acquisition of intra-operative X-ray images. Several researchers explored the

fiducial-based tracking for C-arm assisted navigation [Sie11]. Typical fiducial markers are spherical beads,

straight lines and ellipses. They are recommended due to their projection invariance. An overview of marker-based methods is presented below.

Calibration ring. One of the first approaches has been proposed by Navab et al. [NBHM*96] for interventional angiography. Their work aims at recovering the geometry of X-ray imaging system from each individual X-ray image in order to reconstruct arterial vessels from multiple planar radiographs. They designed a calibration ring which has to be placed around the patient's head as illustrated in Figure 4.10. The ring consists on uniquely identifiable markers (613 steel balls) positioned on a cylinder. The steel balls (ø1.2 mm) have higher absorption coefficient compared to human bones, and can therefore be easily detected and segmented in the X-ray images. However, due to the high number of the fiducials, anatomical structures are partially obscured on the X-ray images. The authors proposed therefore to restrict the placement of the steel balls to a width of 8.5 mm such that only a small fraction of the X-ray image get obscured.



(a) Calibration Ring: perspective view(left), top view(right), and a typical projection view (bottom)

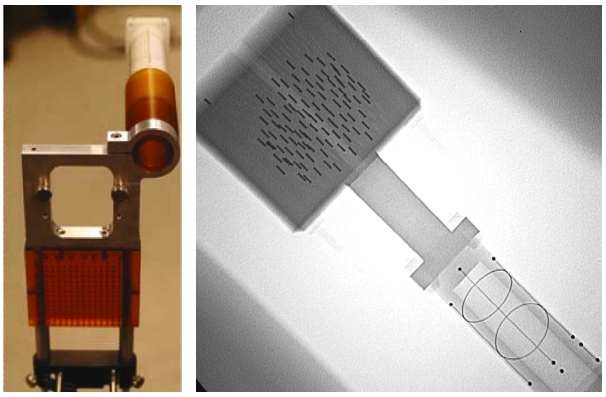
(b) A radiograph which includes the region of interest

Figure 4.10.: The calibration ring used for C-arm pose estimation (a) and an appropriate radiograph (b) [NBHM*96].

In order to determine the C-arm pose, the markers are detected/segmented in the X-ray images by applying a Canny edge detector, clustering curved segments belonging to the same circles, and then estimating the center and radius of markers by applying a circle fitting algorithm. After that, the perspective projection of front part and rear part of the ring onto the X-ray image are separated in upper and lower coded bands. In each band the number of circles are counted for decoding purpose. The found codes are then used to match the projection sequence to the 3D model, which allows the calculation of the projection matrix.

The proposed calibration method has been tested based on 11 X-ray images acquired over limited translational and rotational ranges of ± 10 mm and of 45°, respectively. The reported translational error was up to 0.7 mm, and the rotation uncertainty up to 0.22°. The achieved accuracy is not sufficient for the problem studied in this thesis. Moreover, the limitation of the scan range (due to the high number of the fiducials) imposes additional constraints on the intra-operative navigation.

Fluoroscope tracking (FTRAC). Jain et al. proposed a fluoroscope tracking (FTRAC) fiducial for prostate brachytherapy [JMZ*05]. This device, shown in Figure 4.11, can be mounted over the seed insertion needle template. The designed FTRAC does not only contain spherical beads but also straight lines and ellipses which are invariant to projection. Lines and ellipses can be easily segmented, but with a relatively high uncertainty. A such combination was made in order to easily estimate the C-arm pose from a single X-ray image without degrading the accuracy. The pose estimation is done by aligning the the 3D geometrical model of the FTRAC device to its 2D projection on the detector plane. Therefore, the fiducial have to be segmented in the X-ray images. This task is performed semi- (not full-) automatically, which is inappropriate for clinical application.



(a) (b)

Figure 4.11.: The FTRAC fiducial mounted over the seed insertion needle template using a mechanical connector (a). An X-ray image of the FTRAC fiducial including ellipses, ball bearings and lines (b).

The proposed device has been clinically validated in [JDI*07]. The authors reported an error of 0.56 $\pm$0.33 mm in translations and 0.33°$\pm$0.21° in rotations.

Randomly distributed features. Jain and Fichtinger [JF06] have developed a framework to use randomly distributed feature points present in the imaging volume to determine the C-arm pose. If such features are not present in the scanned region, they can be placed by randomly sticking e.g. beads in the image space. After localizing the feature points in the X-ray images, a non-linear high-dimensional optimization algorithm is employed to find the projective transformation that converts the 3D points to their projections on the X-ray images. At least three X-ray images are required. The authors reported a tracking accuracy of 0.9° and a TRE of 0.8 mm, which is insufficient for skull-base surgery. Furthermore, this method is thought for applications with no need for real-time tracking.

Coplanar ellipses. Burkhardt et al. [BJF07] proposed to use circles or ellipses as fiducial markers since a single circle or ellipse allows the estimation of five degrees of freedom. This idea has been extended by Ayad et al. [ALD*10] to use a set of coplanar ellipses, e.g three ellipses as shown in Figure 4.12, to determine all DoFs, i.e. the C-arm pose. The achieved pose estimation accuracy was 0.9 mm and 0.5° for translational and rotational parameters, respectively. The evaluation has been performed based only

on simulated data. A shortcoming of this method is that ellipses may disturb the diagnosis because they obscure anatomical structures in the projection images as illustrated in Figure 4.12.

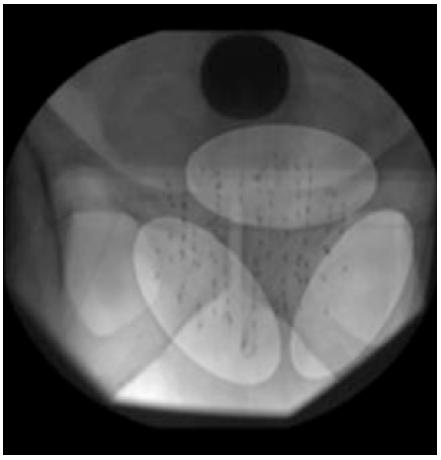


Figure 4.12.: X-ray image which includes ellipse projections [ALD*10].

Marker pattern. Steger et al. [SH12, SDW14] used a specifically designed marker pattern to estimate the C-arm pose for navigated bronchoscopy. They developed an acrylic marker plate which can be easily fixed on any patient table (Figure 4.13). It contains steel markers which are well-visible in the X-ray images. These fiducials define a pattern which allow the unambiguous estimation of the C-arm pose relative to the marker plate, i.e. the patient table, by employing the projective invariant cross-ratio. The pose estimation consists of these four steps:

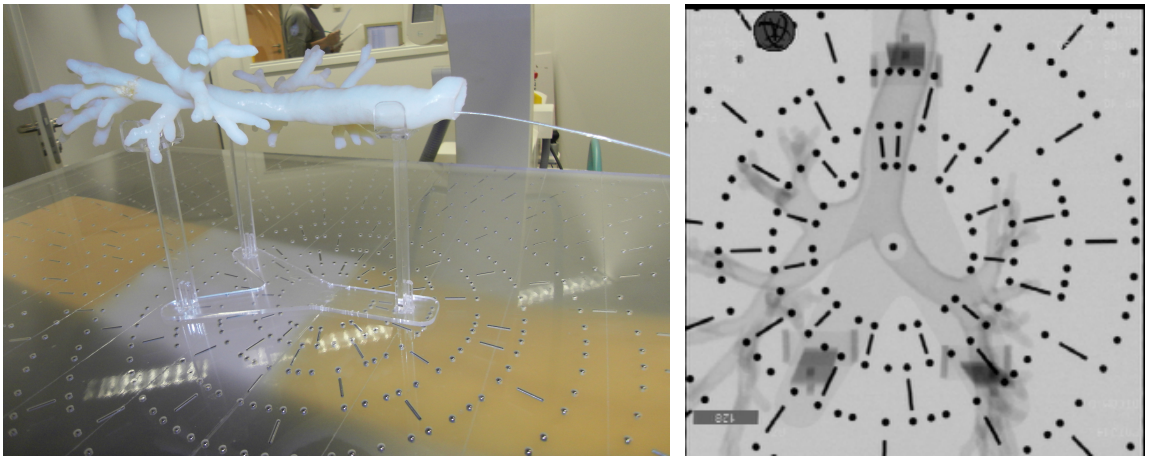
1. detect spheres and lines in the X-ray image
2. localize the center point of the plate defined as the intersection of all lines in the pattern
3. map the detected lines and circles to the corresponding angles and points in the pattern using cross-ratio values and stick positions
4. compute the projection matrix

This method was evaluated using both DRRs and real X-ray images acquired with a C-arm. The reported target registration accuracy (TRE) is 0.33 $\pm$0.2 mm with a success rate of 97.2% over all test images without overlay. In case of overlaying the TRE was 0.47 $\pm$0.35 mm with a success rate of 97.2%. The achieved accuracy is not sufficient for the lateral skull surgery. Moreover, the use of the plate would impose additional contains to the clinical work-flow since the patient's skull has to be fixed to the plate/table.

Common drawbacks. Several marker-based C-arm pose estimation methods exist, but all of them suffer at least from one of the following drawbacks:

- not suitable or cannot be easily integrated into common operation rooms
- the used calibration phantom is not appropriate for temporal bone surgery in term of size and capture range for fluoroscopic images
- the achieved target registration error exceeds the required accuracy of 0.5 mm

The accuracy of feature based methods depends on the localization accuracy of the fiducial marker. Therefore, the focus is set below on marker segmentation methods.



(a) (b)

Figure 4.13.: Airway model on marker plate in C-arm imaging field(a) and a corresponding fluoroscopy (b) [SDW14].

X-Ray Marker Localization. A prerequisite for good pose estimation results is the accurate localization of the fiducials in CTs as well as in X-ray images [FW01]. Performing this step manually is time consuming and subjective, whereas an automatic localization with high accuracy is very challenging in case of X-ray images. The main challenges thereby are mainly the following:

- the marker shape looks different depending on the projection angle due to the perspective projection. For instance, circles are typically projected as ellipses on the X-ray images.
- the size of the projected marker shape varies relative to its distance from the detector and the X-ray source: they appear smaller when they get close to the detector
- the markers have different intensities: depending on the background they appear darker in the raw X-ray image if they are e.g. aligned with bone, and brighter if they are not obscured by other structures.
- low contrast between the markers and the overlaying anatomies or background
- markers are partially occluded or obscured by other objects, like patient anatomies or surgical instruments, in the X-ray image

Usually, small spherical markers, like solid gold spheres, are used for intra-cranial and spinal applications. Such markers have high X-ray mass attenuation coefficients, thus they appear as dark regions in the X-ray images and can therefore be easily detected. However, they have the shortcoming of causing typical streak artifacts in reconstructed CT volumes. Furthermore, such markers might disturb the physician if they occlude important anatomical structures on the X-ray images. Using other materials with low attenuation coefficients causes less reconstruction artifacts, however such markers are more difficult to localize in the X-ray images due to their low contrast. This problem is evident especially if the projected markers are obscured by other overlapping objects.

Several algorithms have been proposed to localize fiducial markers in X-ray images, and all of them follow effectively this pipeline:

1. pre-process the X-ray image, e.g with a gamma correction, in order to enhance the contrast between the fiducials and the surrounding regions (background)
2. find potential marker candidates

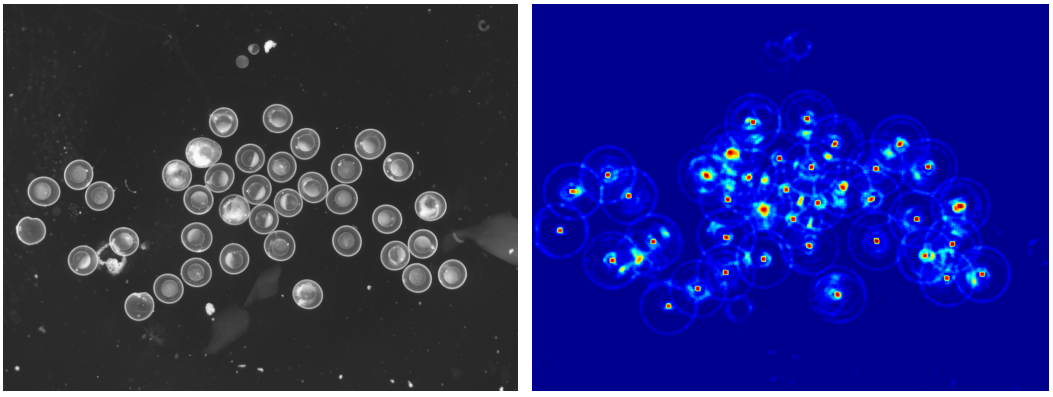
3. determine the centers and radii of the marker candidates

The second step results typically in false positives, however this is not critical since they can be filtered in the optimization step of the calculation of the C-arm pose. The majority of proposed methods focus on the third step because it is decisive for the localization accuracy.

A first alternative to determine the center of a marker candidate is to calculate the intensity-weighted centroid of the corresponding region in the X-ray image [vEPR02]. This method is simple and computationally efficient, but it is not accurate if the markers are partially obscured by other structures in the X-ray image. Obviously, the obscured parts appear darker compared to the non-obscured parts, which results in shifting the centroid in direction of the dark region.

In [YMW09] the localization succeeds by performing an edge detection with Canny edge in a predefined region of interest, followed by circle estimation using the Random Sample Consensus (RANSAC). A similar approach is presented in [Ste14] where the marker centers are located by finding partial arcs which may describe the same circles, and then calculate the centers and radii of these circles. These two methods can easily stuck in local optimum, especially if the marker has a low contrast such that the marker edges can not be detected.

The Hough transform provides a mathematical framework to detect circles. An accumulator which is a matrix filled by the number of border points of circles with a given radius is created. Then the maximum values in this accumulator are extracted: They are supposed to correspond to the circle center points (see Figure 4.14). An optimized Hough transform segmentation was presented in [KAPS09] to segment spherical objects which are occluded by other structures: The Hough transform was extended by a voting framework based on the pixel location and intensity in order to achieve robustness against noise and occlusions. The main drawback of this method is the possible stuck in local optima.



(a) (b)

Figure 4.14.: Hough segmentation example on a zebrafish embryo image (a) and the corresponding accumulator (b) [KAPS09].

Another alternative for marker localization is to perform a template matching using normalized cross-correlation. A marker is located by determining the position where the correlation is maximum. Hamming et al. [HDIS09] presented an approach which finds the markers by setting an appropriate intensity threshold. A pattern matching is then performed based on prior knowledge about the marker size and shape

to localize the fiducials. This method assumes that the marker region has uniform low intensity, which is obviously not satisfied in case of X-ray images.

A Radial Ray based segmentation method was presented by Steger et al. [SEE11]. It was suggested to extract contour of lymph nodes in CT images by minimizing an energy function which depends on gradients, intensities and shape priors. Starting from a seed point within the marker, defined e.g. by the user, rays are casted radially in all possible directions and each ray is uniformly divided into bins as illustrated in Figure 4.15. The method tries then to find the optimal bin (for each ray) which should be closest to the correct marker contour. In this context a cost function which has to be minimized is defined. It consists of a data term depending on the image, and a shape preserving term which incorporates prior knowledge about the shape to be segmented.

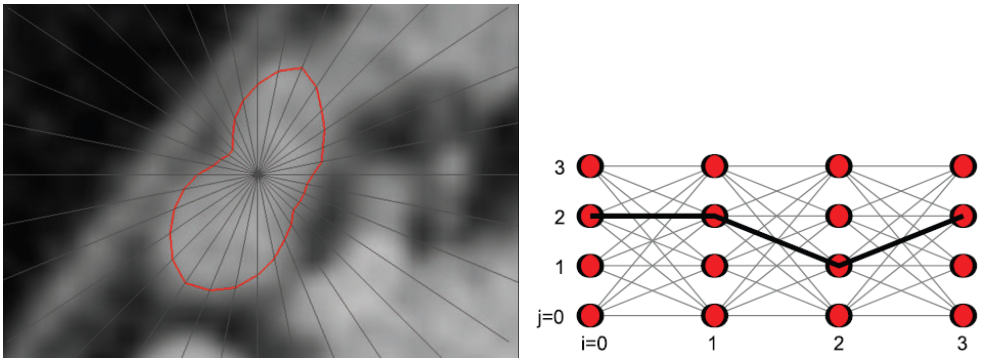


Figure 4.15.: Illustration of the Radial Ray based segmentation method [SEE11].

The robustness and/or accuracy of the listed marker localization techniques are questioned in case of lateral skull surgery. A robust method should be developed in order to accurately localize the marker center points in X-ray images, even if the fiducial are obscured by other structures, or the contrast to the surrounding region is very low. If this step is mastered, the high accuracy of the C-arm pose estimate is guaranteed since the rest of the pipeline is relatively simple to solve with high accuracy.

4.2.2. Concept of pose estimation

In order to align the actual position of the drilling instrument to the patient's anatomy and accordingly to the planned drilling paths, it is mandatory to establish the spatial relationship between the pre-operative CT volume and the intra-operatively acquired X-ray images. The pose estimation is performed by employing a hybrid method which combines the high robustness of feature-based registration and high accuracy of intensity-based registration. It is well known that the feature-based registration is very robust since it can always successfully estimate the C-arm pose if enough features are found. However, its accuracy is inferior compared to the intensity-based registration [PDBF12]. On the other hand, the intensity-based approaches may easily stuck in local optima. In order to overcome these drawbacks an extrinsic registration is employed to compute a transformation which is very close to the correct alignment, then provide this later as initialization for the intensity-based registration.

The proposed pose estimation approach, depicted in Figure 4.16 includes the following steps:

1. localize the fiducials in the pre-operative CT volume and compute their center points
2. localize the fiducials in the intra-operative X-ray image and compute their center points
3. perform a point-based registration of the intra-operative X-ray image to the pre-operative CT volume
4. perform an intensity-based registration after initializing it with the transformation computed in the third step.

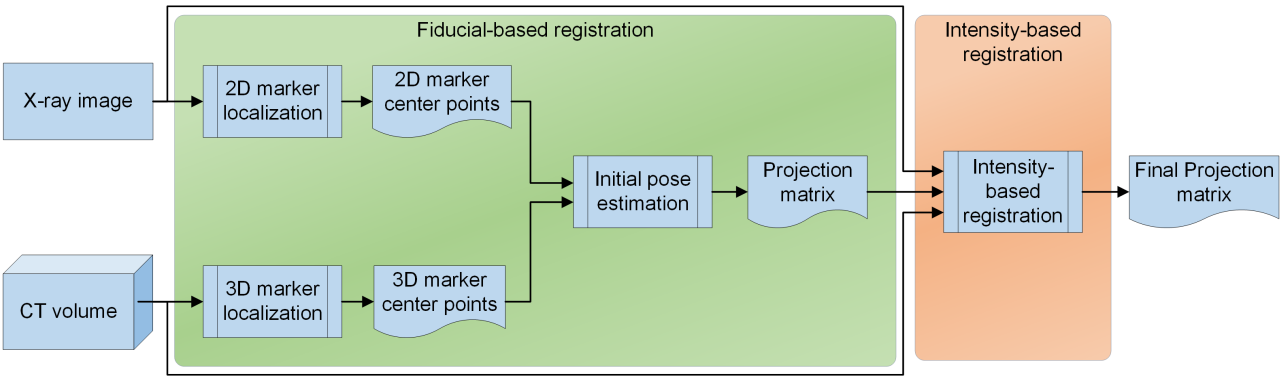


Figure 4.16.: Pipeline of the hybrid registration applied to align an intra-operative X-ray image to a pre-operative CT volume.

Accurate localization of the marker center points in the pre-operative CT as well as in the intra-operatively acquired X-ray images is a prerequisite for an accurate C-arm pose estimation. The localization of fiducial markers in CT volumes is relatively easy, especially when the marker form, size, and the corresponding attenuation coefficient are known (see Section 4.2.3). In contrast, the localization of fiducials in X-ray images is more challenging due to projective form distortion and structure overlapping which would obscure the fiducial markers in the X-ray images. In order to overcome this problem, a novel approach is presented in Section 4.2.4. After determining the center points of the fiducial markers in the pre- and intra-operative images, an initial pose estimate is computed as described in Section 4.2.5. A fine-tuning is then performed by employing an intensity-based method as detailed in Section 4.2.6.

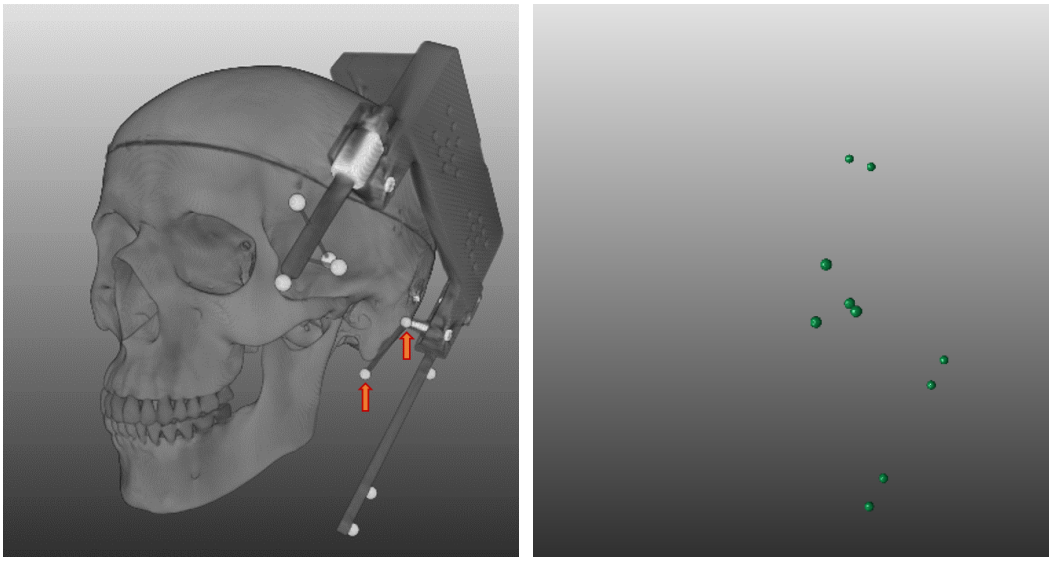
4.2.3. Localization of markers in CTs

The pre-operative CTs are typically acquired with spiral CT scanners, which means that these data sets are Hounsfield-calibrated. This simplifies the marker detection since the density range (attenuation coefficient) of the fiducials is known. Accordingly, the localization of the marker center points is performed as following:

1. a 3D mask (binary image) is generated based of the expected marker attenuation coefficient: if the voxel value lies within the expected density range, its corresponding mask value is set to one, else it is set to zero
2. a connected components labeling [ST88] is performed to extract the different spatially unconnected objects present in the mask
3. filter the segmented objects in order to remove the false positives by employing prior knowledge about the marker shape and size. This results in a list of potential marker candidates

4. localize the center points of the marker candidates by employing an intensity-weighted localization [WMFM96]. Thereby ideal spheres with reference marker proprietaries (size, shape, intensity) are defined as templates
5. asses the localization process by comparing the found marker configuration with the known reference configuration, namely the distances between the marker center points

At this level a set of the marker center points is created (see Figure 4.17).



(a) Rendering of a skull phantom to which a base plate is fixed

(b) Final marker candidates after filtering the false positives

Figure 4.17.: Example showing a pre-operative CT image (a) and the segmented markers (b). Two additional markers smaller as the fiducials used for the registration (red arrows) are placed to check the adequacy of the 3D marker detection algorithm.

4.2.4. Localization of markers in X-ray images

The localization of markers in the X-ray images consists of these two main stages:

1. detect potential marker candidates which are rough estimations of the marker regions
2. determine the center points of the found potential candidates

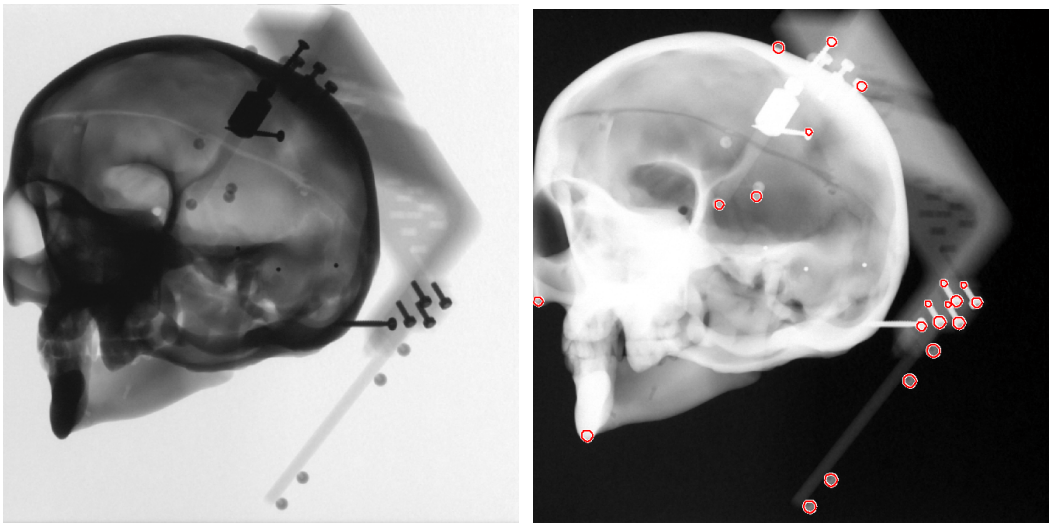
These two steps are described with more details in this section.

Detection of marker candidates

Conventional methods, such as the watershed segmentation algorithm [CBNC09] and Hough circle detection approach [KAPS09], are not robust enough to extract potential marker candidates in X-ray images due to the variance in the appearance of marker projections on the X-ray images. For illustration , the Hough circle detector fail to localize most of the fiducial markers in the X-ray image as depicted in Figure 4.18. Furthermore, many false positives (over 66 %) are found, which would hamper the registration process. In order to improve the marker detection, additional prior knowledge about the used fiducials is employed as described below.

Initially, rough estimation of the marker regions are found by following these steps:

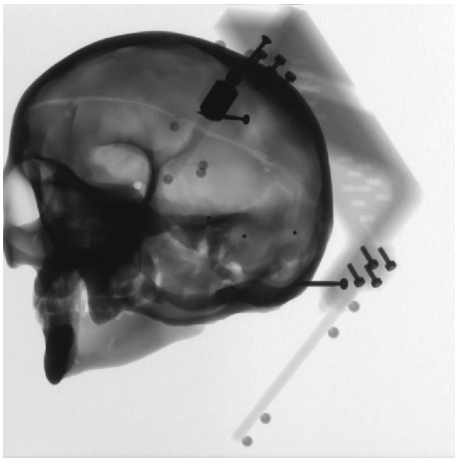
1. Image binarization: It aims to separate the objects (structures present in the X-ray image) from each other as well as from the background. The markers are usually characterized by high gradients on the contours. Therefore, a gradient image is derived from the acquired X-ray image (see Figure 4.19(a,b)). Then, a binary image is computed by applying an adaptive thresholding (minimum cross entropy thresholding [LL93]) to the generated gradient image. As illustrated in Figure 4.19(c), the resulting binary image highlights the contours of the scanned objects.
2. Connected Components Labeling: A label is assigned to each connected region of the foreground pixels by applying the Connected Components Labeling approach proposed in [ST88]. This step results in a list of all spatially unconnected objects (labels) as illustrated in Figure 4.19(d).
3. Removing the false positives: The labels corresponding to the fiducial markers have to be extracted from the previous list. Therefore, prior knowledge is exploited to detect the false positives, namely
   - size (number of pixels): Given that the marker size as well as the source-detector-distance are known, it is possible to estimate the size (in pixels) of the marker projection region on the X-ray image.
   - roundness: The markers have a spherical shape and are therefore projected on the detector plane as ellipses or circles. The roundness of a given label is expressed as the ratio between the label size to the equivalent radius. This ratio is expected to be equal to π in the ideal case.



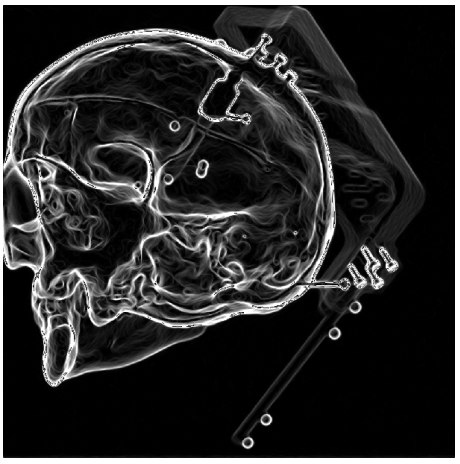
(a) Acquired X-ray image (b) Hough circle detection

Figure 4.18.: Detection of potential fiducial marker candidates using Hough circle detector. Not all marker can be detected and several false positives are found.

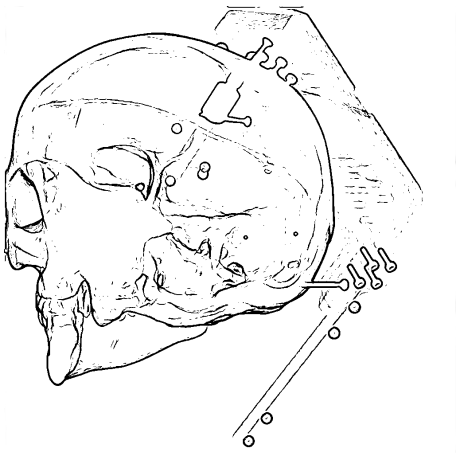
The found potential marker candidates define rough estimations of the marker projection regions. They are set as input for the particle path segmentation (described in the next section) which aims to accurately estimate the corresponding center points. In this stage it is not possible to guarantee that all false positives are detected. However, this is not critical in case of few false positives since they would be excluded in the next step, namely the optimization (calculation) of the C-arm pose.



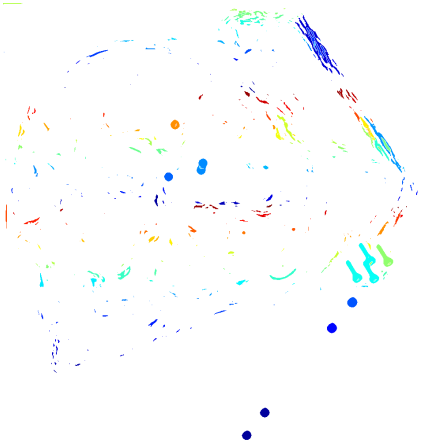
(a) Acquired X-ray image



(b) Gradient image



(c) Binary image



(d) Connected components (labels)

(e) Filtered marker candidates

Figure 4.19.: Detection of potential fiducial marker candidates in X-ray images (a): first a gradient image is calculated (b) and binarized (c), then the connected components are labeled (d). The found objects (labels) are then filtered based on prior knowledge about the marker shape and size in order to remove the false positives. The remaining labels represent the final marker candidates (e).

Particle Path Segmentation (PPS)

This thesis contributes a novel method for segmentation of spherical markers in an X-ray image, which we call particle path segmentation (PPS) [EBS14]. It aims to find the most stable marker contour according to criteria based on marker shape, intensity gradients, magnitude and orientation. Given that spherical markers show an elliptical shape on radiographs, a typical gradient evolution can be computed. This is used to decide the adequacy of the input pixel values and their contributions to update the segmentation of the marker contour. This concept should solve the overlapping problem and offer more robustness in presence of noise. An example illustrating such problems is shown in Figure 4.20.

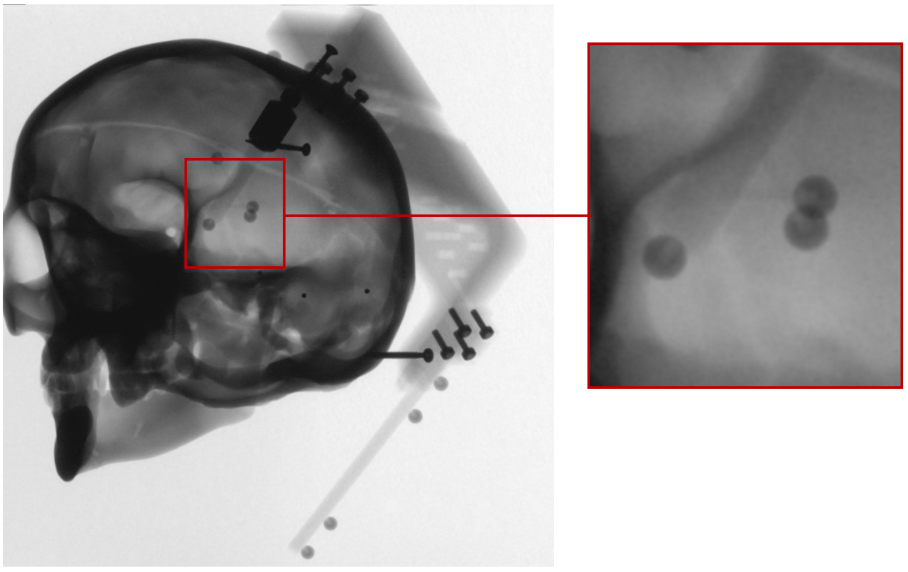


Figure 4.20.: An X-ray image (left) and a zoom on three fiducial markers (right). This example illustrates the overlapping problem which may hamper the accurate localization of the marker center points.

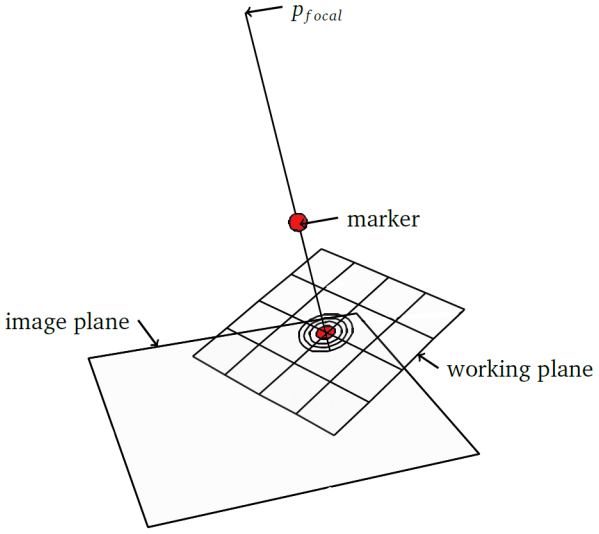


Figure 4.21.: Representation of the distortion free working plane relative to the detector plane. The marker to be localized and its expected projection area are red colored.

It is to be noted that through the perspective projection a spherical shape would be projected as an ellipse so that the sphere center does not anymore coincide with the center of its projected shape. Thus, it is desirable to perform the segmentation in another working space where this geometrical distortion effect is eliminated. A possible plane where projected spheres has circular shapes is the plane containing the center candidate (P_{cc}) and orthogonal to the ray passing through this center point as illustrated in Figure 4.21. Given that the configuration of the used C-arm is known, namely the relative position between the focal point (P_{focal}) and the flat panel detector, it is possible to derive the transformation between the detector plane and the working plane. After performing the segmentation process in the distortion-free working plane, the found center point is projected back onto the detector plane. In the rest of this section the particle path segmentation is done in the working plane, which means that the PPS have to look for circular shapes.

The particle path segmentation (PPS) aims to segment circular objects by rotating a particle around an initial center candidate and evaluating its silhouette based on the image appearance. The method tries to find the most stable shape describing the contour of the projected marker. Therefore, the PPS is formulated as a system consisting of three interactive components as illustrated in Figure 4.22:

- **Particle:** a mass point associated with a position x_t (= $x(t)$) and a velocity vector $\vec{v}(t)$ at time t. It is rotating around a given center candidate P_{cc}.
- **Silhouette:** an approximation of the potential circular marker contour. Its center is the current center candidate P_{cc}. The silhouette itself is evolving over time and can be influenced (expanded or compressed) by the particle if a more stable contour is found.
- **Score system:** a system rating the reliability of the actual particle path. It decides about the evolution of the silhouette and the particle course.

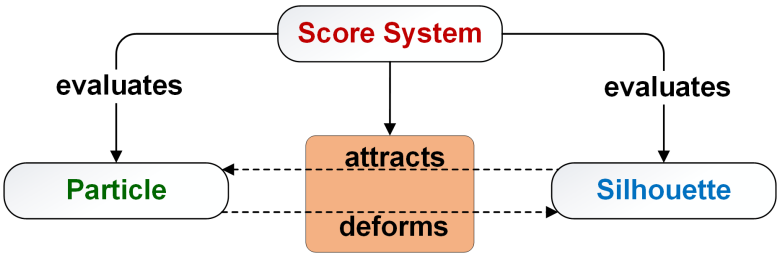


Figure 4.22.: The interaction between the components of the particle path segmentation.

These three main components are described in the next pages, but before that we introduce two simple examples which aim to demonstrate why these components are required.

- Example 1: Let the potential field, computed based on the gradient vector map of the input marker region, have a circular appearance as illustrated in Figure 4.23(a). Let the particle starts rotating somewhere near the center of this region with constant (radial and tangential) velocity (>0). At the beginning the particle follows a spiral trajectory as shown in Figure 4.23(b) since the potential field has no influence on the particle, so that the velocity remains constant (>0). As the particle reaches the marker contour, it get attracted by the potential field there and keeps rotating on the ideal circular marker shape as depicted in Figure 4.23(c). In this case the silhouette concept is not needed.
- Example 2: If the marker is obscured by other high-density objects, the potential field would not have a perfect and complete circular shape, e.g as depicted in Figure 4.24(a). In this case the particle would first follow the potential field and then would leave the correct circular marker contour at the

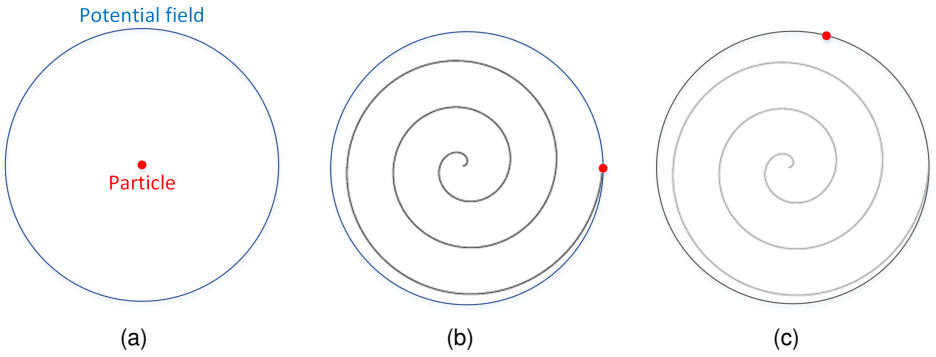


Figure 4.23.: An example illustrating the particle path. For the case shown in (a) the particle follows a spiral trajectory (b) than it is attracted by the potential field and keeps rotating along the ideal marker circular shape (c).

position where the gradient vector is no longer collinear to the radial velocity $\vec{v}_r$, as illustrated in Figure 4.24(b).

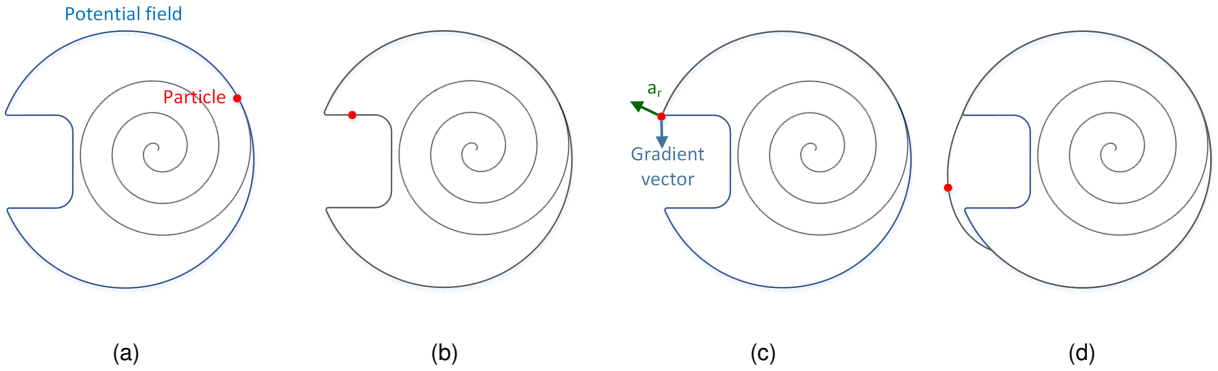


Figure 4.24.: An example illustrating the trajectory of the particle in case of non-ideal circular shape: it follows a spiral trajectory until it gets attracted by the potential field (a). In order to minimize the influence of the potential field at non-circular or missing marker contour, as it is the case in (b), the angle between the radial velocity direction and the gradient vector direction is considered (c). The particle leaves the apparent contour at this position, and may be attracted later by the potential field as shown in (d).

In order to overcome this problem, the influence of the potential field should be penalized at positions where no reliable marker contour information is present. Therefore, the gradient vector direction is considered. Ideally the radial velocity direction and the gradient vector direction on the marker contour should be similar. If this condition is not fulfilled, the potential field should have no influence on the particle course. In this case the particle would be attracted by the potential field only till reaching the position where the gradient vector direction and the radial velocity direction get divergent as depicted in Figure 4.24(c). At this position the particle would be influenced only by the radial velocity and moves therefore away from the ideal marker contour. Sometime later, the particle might be attracted again by the potential field to lie on the marker contour, however it would never follow the circular marker shape over the complete marker contour as illustrated in Figure 4.24(d).

A second mechanism, called here silhouette, is therefore proposed in order to not lose the particle at positions with no reliable gradient vectors. The silhouette incorporates prior knowledge about the marker shape and aims to stabilize the particle trajectory until lying on the ideal marker contour. As depicted in Figure 4.25(b), the particle expands the silhouette with some latency. The silhouette keeps expanding until reaching the expected marker contour as illustrated in Figure 4.25(c). If a reliable contour is found, the particle can not expand the silhouette any more. In contrast, the particle would be attracted by the silhouette to lie on the ideal marker contour as depicted in Figure 4.25(d).

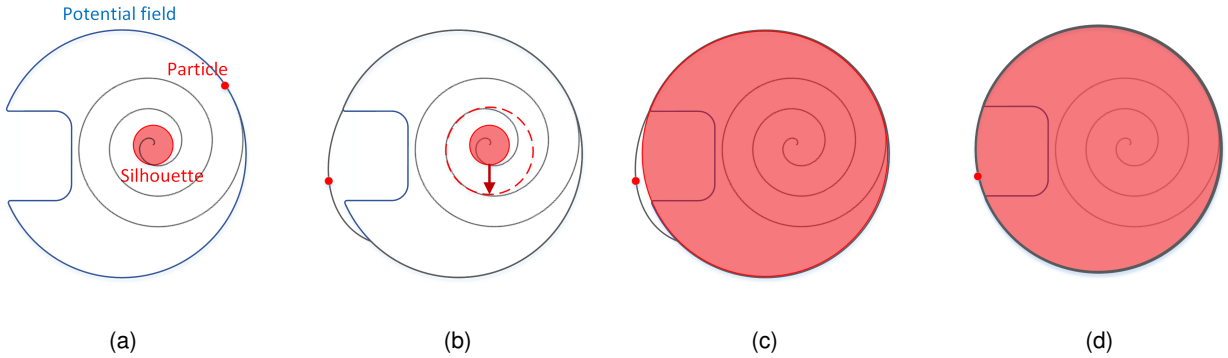


Figure 4.25.: An example illustrating the particle path after employing the silhouette concept: it follows a spiral trajectory (a) and expands the silhouette (b). When the silhouette reaches a stable contour (c), it stops expanding an attract the particle back (d).

Particle. At the beginning of the segmentation process the particle (mass point) as well as the center candidate P_{cc} are placed at the center of the window limiting the input marker candidate region. The particle has at time t the position x_t (= $x(t)$) and the velocity $\vec{v}(t)$. The velocity $\vec{v}$ has two components as illustrated in Figure 4.26:

- tangential (cross-radial) velocity $\vec{v_t}$ which is perpendicular to the ray direction $\overrightarrow{P_{cc}\,x_t}$. It serves to keep the particle in rotation. Its magnitude v_t is proportional to the distance between x_t and P_{cc}. It is to be noted that $\vec{v_t}(t)$ is defined in the working plane, thus

$$\vec{v_t}(t) = \omega \cdot \|x_t - P_{cc}\| \cdot \frac{(P_{focal} - P_{cc}) \times (x_t - P_{cc})}{\|(P_{focal} - P_{cc}) \times (x(t) - P_{cc})\|} \tag{4.1}$$

  where ω is the angular velocity and P_{focal} is the position of the focal point.
- radial velocity $\vec{v_r}$ which is collinear to $\overrightarrow{P_{cc}\,x_t}$. This radial part decides if the particle gets closer to the center candidate P_{cc} or moves away from it. At the beginning its magnitude v_r remains constant (>0) if the particle is not influenced by the silhouette or the image gradient field such that the particle follows a spiral trajectory. Unlike the tangential velocity which remains constant, the radial velocity evolves over the time. If its magnitude is negative, the particle moves inwards toward P_{cc}, otherwise the particle is pushed away from P_{cc} if the said magnitude is positive.

The external forces applied by the potential field as well as by the silhouette to the particle influence only the radial velocity, i.e. only the projections of the said forces on the radial direction are considered. Consequently, the angular velocity remains unchangeable. In contrast, the distance between the particle and P_{cc} evolves over time until a reliable marker contour is found.

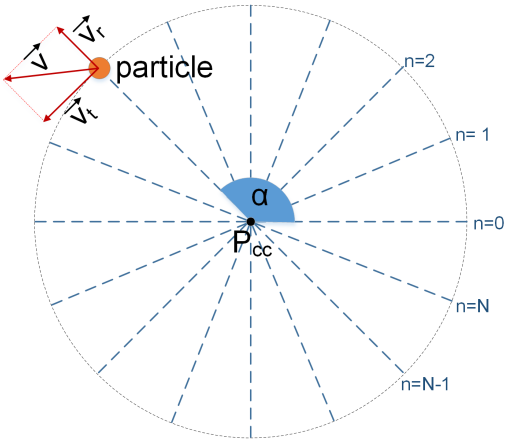


Figure 4.26.: Illustration of the silhouette structure: it is represented as a point cloud which consists of N points.

Silhouette. The silhouette s is a point cloud describing the current approximation of the segmented marker shape as illustrated in Figure 4.26. It consists of N points s_n and is described as following

$$s_n = P_{cc} + r_n \begin{bmatrix} cos(\alpha_n) \\ sin(\alpha_n) \end{bmatrix} \tag{4.2}$$

where r_n is the distance to the center candidate P_{cc}, and

$$\alpha_n = n \cdot \frac{2\pi}{N} \tag{4.3}$$

The evolution of the silhouette shape is performed based on this equation

$$s_i = s_i + f_s \cdot (x_t - s_i) \tag{4.4}$$

with s_i ($\in$ s) the nearest silhouette point to the actual particle position x_t. The term f_s is a flexibility coefficient ($\in$]0, 1]) which defines how much the silhouette should adjust towards each new computed particle position. At each change of the silhouette shape, the position of the center candidate P_{cc} is updated to coincide with the actual silhouette mass center point.

Starting at a given seed point as an initial center candidate, the particle path should converge against a reliable circular silhouette shape. Accordingly, the following constrain is defined to let the distance of all silhouette points from the silhouette center be the same:

$$r_n = \lambda r_n + (1 - \lambda)\bar{r} \tag{4.5}$$

with $\bar{r}$ the mean radius:

$$\bar{r} = \frac{1}{N} \sum_{n=1}^{N} r_n \tag{4.6}$$

The term λ is a flexibility term $\in$]0, 1] which defines allowable deviation level from circular shape: when λ is zero, a circular shape is imposed, whereas an arbitrary shape can be created if λ is equal to one.

At the beginning, λ is set to 1, which allows the individual silhouette points to be adjusted independently of the rest of silhouette points. In this case the silhouette can take any arbitrary shape. With increasing number of iterations the term λ decreases, which imposes the circular silhouette shape since all points would have the same distance from the silhouette center P_{cc}. It is to be noted that the term λ can should not be set to zero in order to allow further evolution of the silhouette.

Score system. The score system aims to stabilize the particle course and deform the silhouette to lie on the marker contour. This contour is usually characterized by locally higher gradients magnitude. Therefore, the gradient magnitude $\|\nabla_x\|$ and the gradient direction (indicated by the angle θ_x) at the particle position x are employed by the score system (see Figure 4.27). In absence of noise and structure overlapping, the gradient direction should be similar to the radial velocity direction, i.e. $\theta_x \approx \beta_x$. The gradient score at a given position x is therefore defined as follows

$$G(x) = (1 - \min(1, \frac{\angle(\beta_x, \theta_x)}{\gamma_{thres}})) \cdot \|\nabla_x\| \tag{4.7}$$

where γ_{thres} is a scaling factor indicating the maximum allowable deviation between the gradient vector direction and the radial velocity direction.

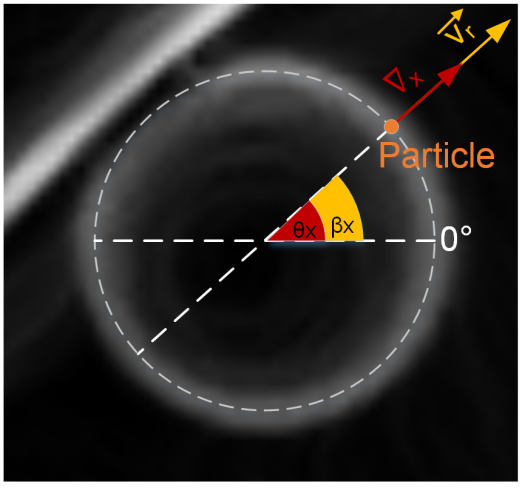


Figure 4.27.: In ideal case the gradient vector ∇_x at each point of the marker boundary is expected to be collinear with the radial velocity $\vec{v}_r$ at the same point, i.e $\theta_x = \beta_x$.

As depicted in Figure 4.28, the decision about the silhouette deformation and the particle guidance is made based on the current particle position (x_t), the position of the nearest silhouette point to it (s_i), the distance to the center candidate P_{cc}, and the corresponding gradient scores ($G(x_t)$ and $G(s_i)$). In situation (A) the particle is "outside" the current silhouette and has a higher gradient score than s_i. In this case the particle is "on the right track" towards the final marker shape and hence the current silhouette shape should be expanded. In contrast, the particle is again outside the silhouette but has in case (C) a lower gradient score than s_i, thus the particle may have passed the marker edge. In this case the silhouette remains unchanged and the particle should move inwards toward the actual silhouette. If the particle is "inside" the silhouette but has higher gradient score than s_i (case B), particle and silhouette are moved towards the center. Last,

if the particle is “inside” but the local gradient score at the particle position is smaller than at s_i (case D), it is more likely that the contour lie in the vicinity of the actual silhouette and the particle is therefore moved outwards towards the current silhouette.

Based on the decision made by the score system as described above, the particle changes its radial velocity as follows:

- if the particle position is more reliable compared to the silhouette, the potential force $F(x_t)$ described below is applied to the particle by influencing its radial velocity $v_r(t)$. Here, we assume that $F(x_t)$ remains constant till the next particle position and we consider only the projection of $F(x_t)$ on the radial direction:

$$v_r(t) = \gamma \cdot F(x_t) \cdot \frac{\overrightarrow{P_{cc}\, x_t}}{\left\| \overrightarrow{P_{cc}\, x_t} \right\|} \tag{4.8}$$

  where γ is a scaling factor serving to scale the influence of the potential force on the particle in order to not move the particle far away from the marker contour
- if the silhouette should attract the particle, this later changes its current position as follows:

$$x_t = x_t + f_p \cdot (s_i - x_t) \tag{4.9}$$

  where f_p is a flexibility coefficient ($\in$]0, 1]) which defines how much the particle should move towards the silhouette point s_i.

Similarly, if the silhouette has to be deformed, it evolves following the Equations 4.4 and 4.5.

Potential field. A potential field is created from an input X-ray image P such that the valleys of this potential field correspond to the location where sharp changes in intensity are recorded, thus the gradients of the input X-ray image are computed:

$$P^{(1)}(u, v) = \|\nabla P(u, v)\| \tag{4.10}$$

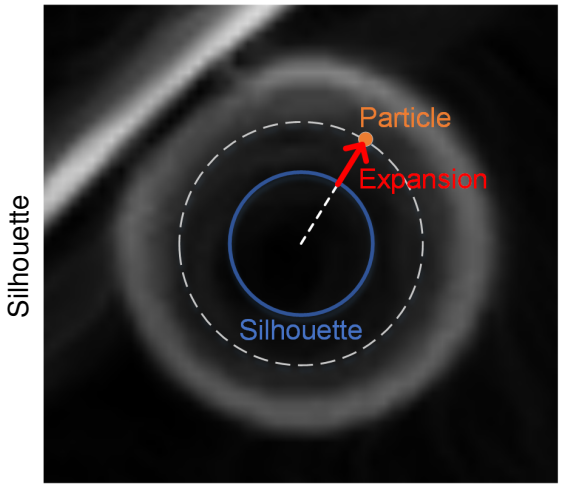
where ∇ is the gradient operator. In this context the potential force field F induced by the potential field is defined as

$$F(u, v) = \beta(u, v) \cdot \nabla P^{(1)}(u, v) \tag{4.11}$$

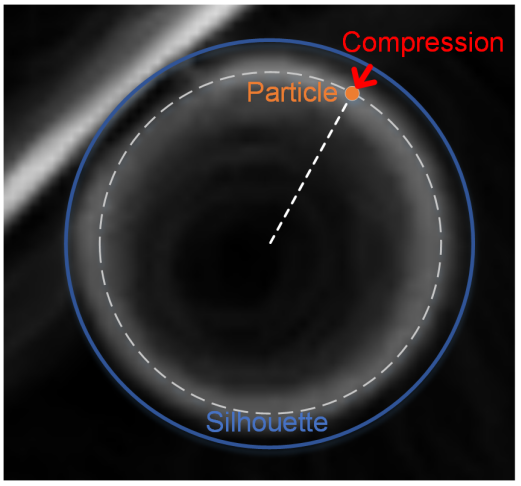
where β is a weighting factor which serves to weight the effect of the potential force field in order to consider it only at positions where the marker contour is expected to be projected. The term β helps to prevent the particle to stuck e.g. in small circular-like shapes which are projected on the same position as the fiducial marker on the X-ray image. The calculation of the weighted potential force field is done as follows:

1. calculate the range of the minimum radius r_{min} and the maximum radius r_{max} of the marker projection on the X-ray image based on the known marker size (radius) and the intrinsic geometric projection parameters.
2. define an annulus to be the region bounded by the concentric circles of the radii r_{min} and r_{max}
3. center the annulus on P_{cc}
4. set the weighting $\beta(u, v)$ to one at positions within the said annulus, otherwise $\beta(u, v)$ is set to zero.

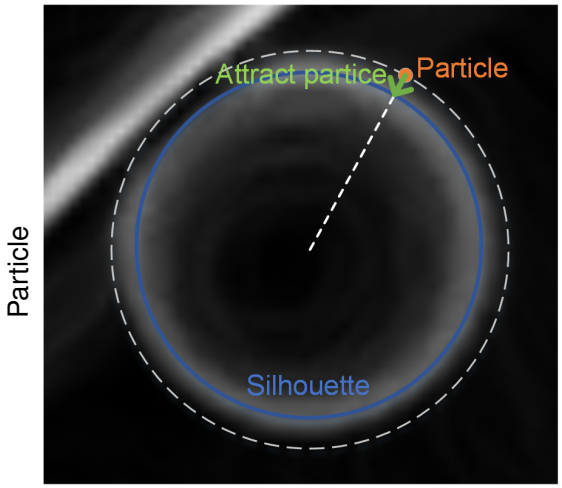
Consequently, the potential forces are considered only at positions where the marker contour is expected to be present in order to achieve more accurate segmentation results.



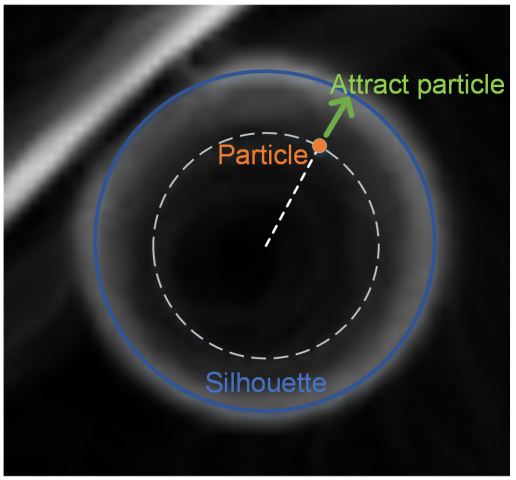
(a) Case A: $G(x_t) > G(s_i)$: expand the silhouette



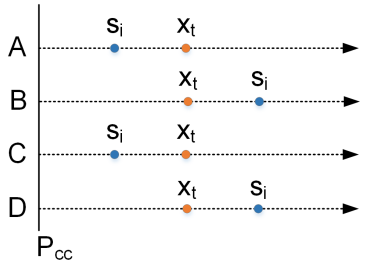
(b) Case B: $G(x_t) > G(s_i)$: compress the silhouette and move the particle towards the center



(c) Case C: $G(x_t) \leq G(s_i)$: move the particle inwards toward the silhouette



(d) Case D: $G(x_t) \leq G(s_i)$: move the particle outwards towards the silhouette



condition	particle	silhouette
$G(x_t) > G(s_i)$	push	expand
$G(x_t) > G(s_i)$	attract	compress
$G(x_t) \leq G(s_i)$	attract	ignore
$G(x_t) \leq G(s_i)$	push	ignore

(e) Possible configurations

Figure 4.28.: Possible configurations of the particle x_t and the nearest silhouette point s_i.

"How it works". The particle plays the role of a scout, searching for gradient fields lying on a more or less circular trajectory. The silhouette follows the particle with some hysteresis and has the tendency to converge to a found circular gradient field. The convergence force for both, particle and silhouette, depends thereby on the length of the circular arc and the gradient strength along the said arc. The particle movement has the tendency to "overshoot" outwards or inwards the silhouette, avoiding thereby convergence to local maxima.

The particle is initialized and starts its movement at the estimated center of the marker, i.e. somewhere more or less close to the center of gravity of the projected marker. Given that no significant gradient is expected there, the centrifugal force is "pushing" the particle along a spiral line outwards. For simplicity let's assume here that both $G(x)$ and $G(s)$ are zero, thus no silhouette movement is initiated yet.

As soon as the particle finds a significant gradient, $G(x)$ becomes stronger than $G(s)$ and the silhouette starts expanding with some latency towards the current particle path (case (A) in figure 4.28). The particle itself shows thereby a tendency to expand further. After some time the silhouette also reaches the location of said first gradient and shows the tendency to rest there. If the particle, which continues moving outside the silhouette border, detects a new gradient outside the first one with a higher significance, the silhouette expansion continues following again case (A). Otherwise if the particle moves in "gradient-free space", i.e. no next gradient layer is detected outside the first one, the radial velocity is reversed and the particle starts moving inwards again (case (B)). If the current silhouette is the most dominant one, the particle starts a smooth oscillation driven by the forces of case (B) and case (D) until particle and silhouette converge to the final circular shape. During this smooth oscillation cases (A) and (C) become alternating active, refining and fine-tuning thereby the shape of the silhouette (and its center) towards the final convergence position. Some examples illustrating the particle path evolution for different background configurations are shown in Figure 4.29 .

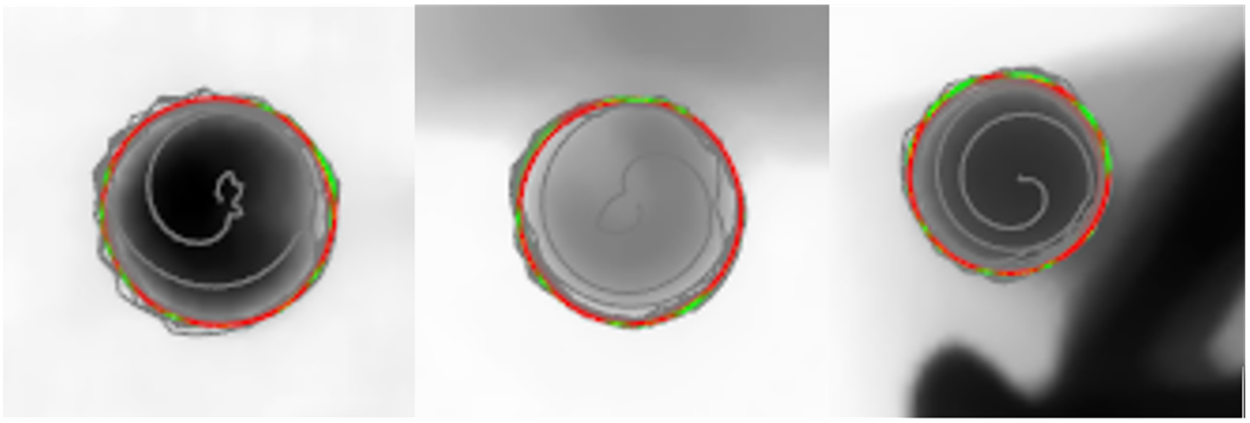


Figure 4.29.: Particle path evolution during the segmentation process. Expansions are marked with red and compressions with green. The color intensity increases with iterations number.

Convergence criteria. The method converges when the distance between the particle and the nearest silhouette point is smaller than a given value $< \epsilon_c$ at each position over a full rotation around P_{cc}. A maximum number of iterations is also defined to stop the PPS if this number is reached before the method converges. This is especially useful in case of false positives where no convergence is expected.

In order to not stuck in local optimum e.g. in case of two concentric circular contours, the following mechanism is employed: given that the PPS has converged as described above, a positive radial velocity is applied to the particle in order to push it away from the silhouette and the center candidate P_{cc}. Two scenarios are then possible:

- if a more reliable marker contour is found, the silhouette starts evolving again to converge towards the new (correct) marker boundary
- if no new (reliable) marker contour is found, the particle is attracted again by the silhouette to converge towards the previously found marker boundary. In this case the segmentation process is terminated

It is to be noted that the weighted potential force field concept described previously serves to ensure that the method convergences towards a marker contour which has a radius in the expected range.

4.2.5. Initial pose estimate

The inputs provided for the initial pose estimate are the marker center points in the CT volume as well in the X-ray images. Several methods which aim to determine the correspondence between two point sets have been proposed and intensively studied in computer vision. In this thesis, the point correspondence and the pose estimation are solved with the SoftPosit method [DDDS04]. It combines the iterative soft assign algorithm and the iterative POSIT method. The first algorithm determines the correspondences between the 3D and the 2D center points, whereas the second computes the C-arm pose by assuming a full-perspective camera model. These algorithms are descried with details in [Fit03, GRL*98, DDDS04].

As result, a 3×4 projection matrix mapping the 3D coordinates to the 2D coordinates is computed. For instance, $\mathbf{H}_1^{(0)}$ and $\mathbf{H}_2^{(0)}$ are calculated for the input X-ray images P_1 and P_2.

4.2.6. Pose fine-tuning

It was shown in Section 4.2.1 that the intensity-based registration tries to find the geometrical projection parameters at which the created DRR is expected to match perfectly with the acquired X-ray image. If this registration process is initialized with the projection parameters computed in previous step ($\mathbf{H}_1^{(0)}$ and $\mathbf{H}_2^{(0)}$), it is expected that the estimated C-arm pose gets closer to the correct pose, i.e. $\mathbf{H}_1$ and $\mathbf{H}_2$.

The common five similarity measures used for 2D-3D registration are the Normalized Cross-Correlation (NCC), Gradient Correlation, Pattern Intensity, Gradient Difference Correlation, and Mutual Information. These metrics have been tested in combination with the three common optimization methods, namely the Powell's method, Downhill simplex algorithm, and genetic algorithm, in this thesis (see section 3.5.4) as well as in several studies in term of accuracy, robustness, and computational efficiency. Experimental results presented in [Ste14, vDH*11, dPAB90] have shown that the combination of Normalized Cross-Correlation with Powell's optimizer achieved the best results for the investigated problem, because the gray levels of the acquired X-ray images and the corresponding DRRs are linearly dependent. The NCC is robust against the expected linear brightness and contrast variation of the created DRRs. The combination of NCC measure method with downhill simplex algorithm should achieve high registration accuracy if the registration process starts in close proximity of the correct pose. Accordingly, the combination of NCC and downhill simplex optimizer are employed in this thesis to perform the intensity-based registration.

Masked registration. The DRRs are created based on the pre-operative CT. This later does not contain any surgical instrument. Consequently, the DRR would not contain the projection of the drill bit. This discrepancy between the acquired X-ray images and the corresponding DRRs would increase the registration uncertainty. In order to overcome this problem, a masked registration is performed. In other words, the projection region of the drilling instrument is not considered in the similarity measure. This succeeds by following these steps as illustrated in Figure 4.30:

1. compute the expected projection area of the drilling instrument based on the planned paths and the initial pose estimate. Assuming that the drill is close to the planned drilling paths and that the initial pose estimate has high accuracy, the estimated projection area would offer a good approximation of the current drill bit position in the X-ray image
2. generate a binary mask whose values are set to zero at the expected drill projection region, and to one at the rest of pixels
3. dilate the binary mask with in order to account for instrument positioning error and initial pose estimation uncertainty. The kernel size is set to 15×15 pixels
4. pixels whose mask value are equal to zero doe not contribute to the similarity measure

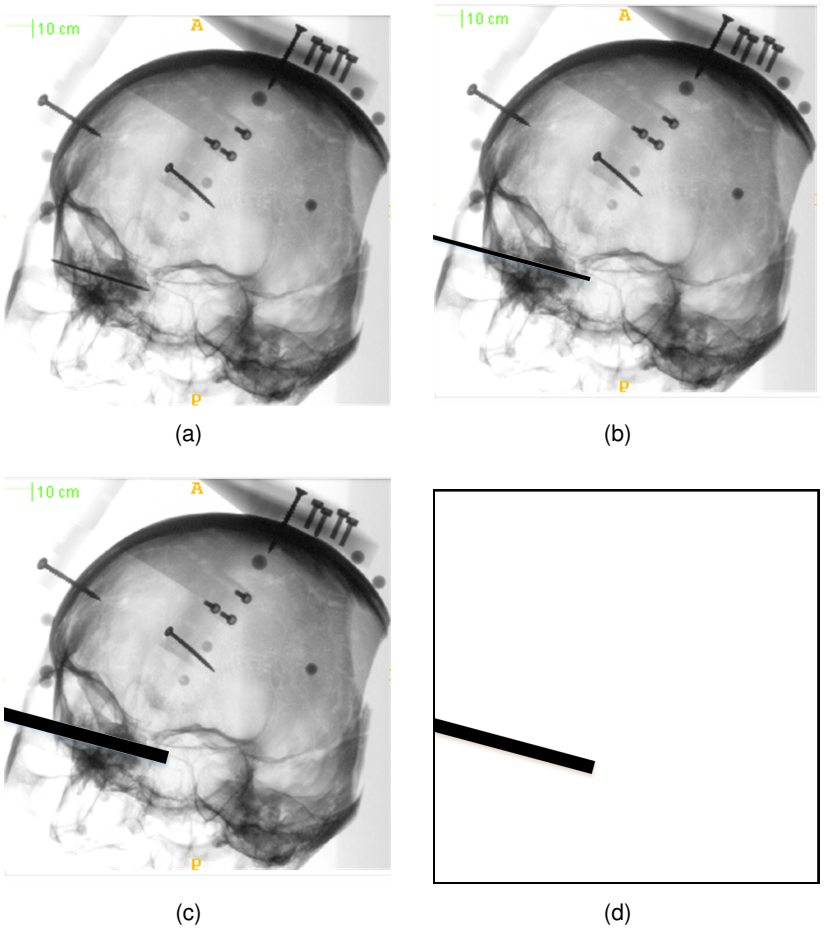


Figure 4.30.: Masked registration: the projection region of the drilling instrument on the X-ray image (a) is computed (b), this region is dilated (c) and used to define a binary mask (d). The black pixel in (d) are not considered in the similarity measure, unlike the white pixels.

4.2.7. Evaluation of the pose estimation

A quantitative analysis based on real X-ray images was not possible due to the missing reference data: even manual segmentation on X-ray images is not accurate enough to be considered as reference segmentation! Therefore, a simulation study was performed by randomly generating a set of 5.000 digitally reconstructed radiographs (DRRs) from a CT volume. Ten fiducial markers made of synthetic ruby were mounted on the designed base plate which were fixed on a skull phantom. Then, a CT volume was acquired with a spiral-CT scanner (SOMATOM Definition AS, Siemens AG). The geometric parameters employed to generate the DRRs were set similar to the configuration of the C-Arm CBCT System described in Section 3.2.1.

Patches containing the marker projection regions were cropped and set as input for two conventional localization methods and the proposed approach:

- radial ray based segmentation (RRS) [SEE11]
- Hough based method (Hough) [CBNC09]
- the proposed approach (PPS)

Fiducial localization error (FLE). A quantitative evaluation was performed by means of fiducial localization error (FLE) which is the Euclidean distance between the localized and the reference marker center. The ground truth is computed based on the (known) projection matrix used to generate the DRRs. As shown in Figure 4.31, the proposed method has the highest accuracy with minimum variation (mean: 0.059 mm; std: 0.062 mm). If the fiducial markers are overlapping with other objects in the X-ray images, the proposed method as well the radial ray based segmentation still perform well, because they implicitly vote the adequacy of the input information (pixel intensities). The Hough based method considers all edges surrounding the marker, including contours of other objects, which leads to wrong results.

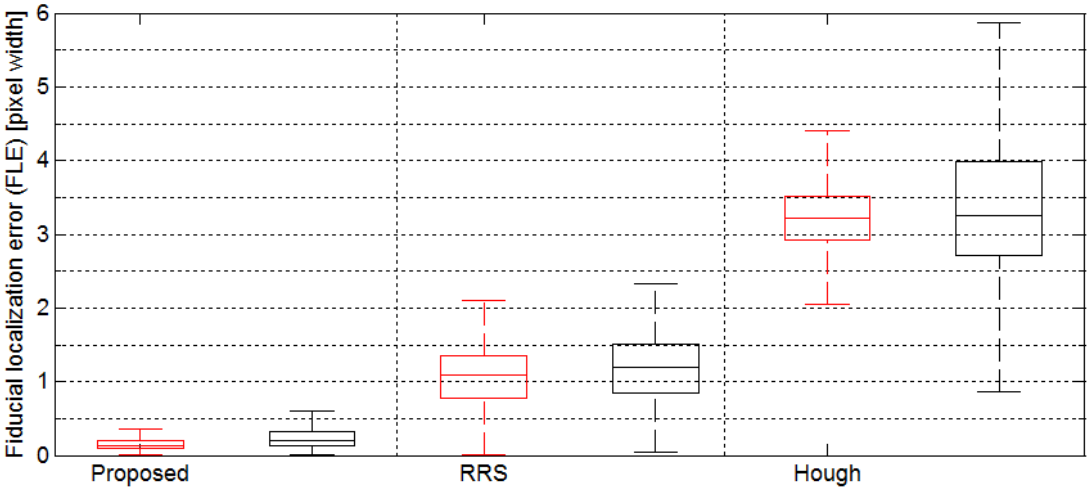
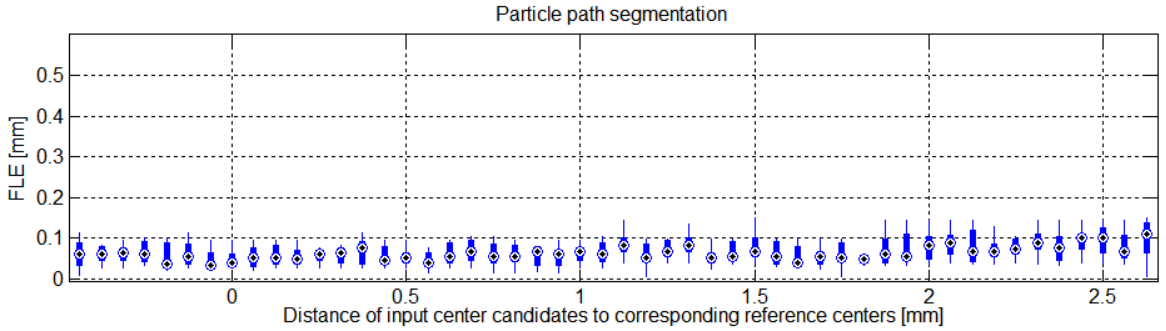
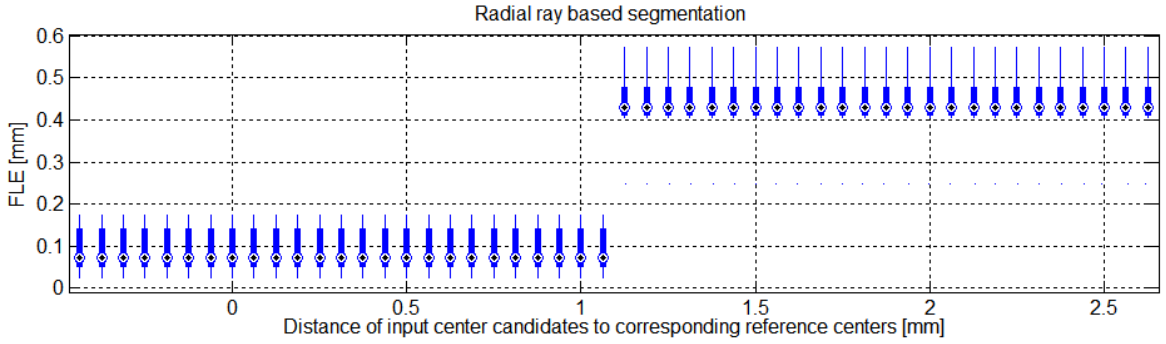


Figure 4.31.: Fiducial Localization errors computed for the case in which the markers are obscured by other objects in the X-ray images (black) and for the case of no overlapping (red). The methods to be compared are the proposed approach, the radial ray based segmentation (RRS), and the Hough method.

Influence of initialization error on the localization accuracy. An initial marker center point is set based on a rough estimation of the marker region. In order to study the influence of the marker center points initialization accuracy on the marker segmentation results, the previous tests are repeated while inducing controlled shifts to the initial marker candidates. For instance, if the shift is equal to zero, the input patch (marker candidate region) is centered at the correct marker center point. As shown in Figure 4.32(b), the radial based segmentation may stuck in local optimum if the shift is relatively high. In contrast, the proposed method always converges against the correct marker center as long as the initialization lies somewhere within the marker contour.



(a) Proposed method (PPS)



(b) Radial ray based segmentation(RRS)

Figure 4.32.: Influence of the initialization uncertainty on the accuracy of the localization of the marker center points.

C-arm pose estimation accuracy. First, we evaluated the influence of the marker localization method on the C-arm pose estimation accuracy based on simulated data. DRRs were created from the mentioned CT volume (described above) and provided as input for the C-arm pose estimator described in Section 4.2. The target registration error (TRE) were computed for each DRR based on the known correct projection parameters. This test was then repeated while changing the method used to compute the marker center points in the X-ray images (here the DRRs). As shown in Figure 4.33, a mean TRE of 0.081 mm (std=0.066 mm) was achieved by employing the proposed PPS method. For comparison, the computed TRE for RRS and Hough was respectively 0.37 (std=0.18) mm and 0.80 (std=0.32), i.e. 5 to 10 times larger.

The target registration error was also evaluated based on real data when the proposed pose estimation method combined with the PPS. Therefore, three small metallic beads were placed on the skull phantom. These fiducials were employed to determine the TRE, but not used in the pose estimation process. The

achieved TRE was 0.167(std: 0.018 mm), i.e. twice larger compared to the simulated study. This is basically due to noise.

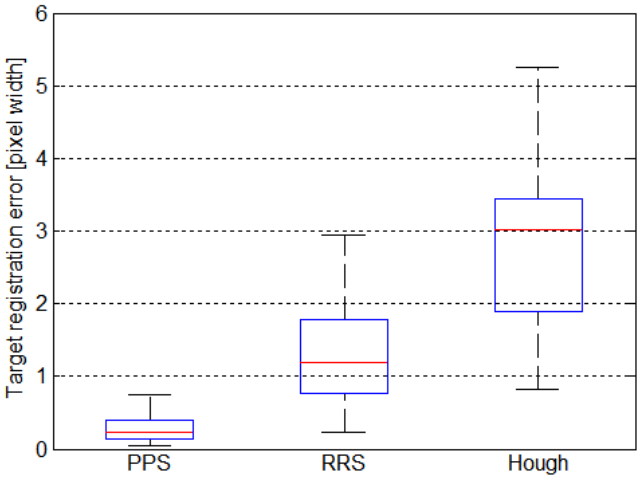


Figure 4.33.: Target registration error for point based registration performed after locating the marker centers in the X-ray images using the proposed method (PPS), the radial ray based segmentation (RRS), and the Hough method.

Computational complexity. The tests were performed on an Intel Core i5-4670 CPU with 3,40 GHz CPU and 16 GBytes RAM. The DRR generation was performed on a GeForce 7600 GTX graphics card. The CT volumes have a dimension of $512 \times 512 \times 512$ voxels, while the DRR image size is set to be 1024×1024 pixels.

The calculation of the marker center using the particle path segmentation (PPS) requires in average 0.011 sec per fiducial and 0.24 sec to localize all markers in an X-ray image. The point-based pose estimation, i.e. the correspondence search, takes in average 0.46 sec.

Table 4.1 summarizes the processing time required to complete the C-arm pose estimation including the pose refinement with intensity-based registration. The mean required time is 5.97 sec. A mean speed advantage of factor 46 was obtained with the GPU implementation compared to CPU. This significant speed up allows the application of the proposed C-arm pose estimation for image-guided navigation in clinical practice.

	CPU	GPU	Speed-up
min	251.3s	5.87s	42.82
max	317.9s	6.39s	49.76
mean	276.47s	5.97s	46.31
std	19.78	0.153	–

Table 4.1.: The execution times needed for the C-arm pose estimation.

4.3. Instrument localization

An essential step in the image-guided navigation is the instrument tracking. The systems of choice for this task are the camera-based optical tracking and electromagnetic tracking (EMT) systems. The main

advantage of optical trackers is the high measurement accuracy and large field of view. However, they require a clear line-of-sight between the tracker and the instrument [KKK13], which can not be satisfied in case of minimal-invasive temporal bone surgery since the tip of the drill instrument is inside the skull. In contrast, EMT systems do not have such a constraint. They localize small sensors (up to 0.3 mm) in a pulsed electromagnetic field with 5 or 6 degrees of freedom. However, this field is susceptible to distortions caused by metallic objects, which would increase the measurement errors in the clinical environment [LAH03]. In effect, the reported target registration error (TRE) of the mentioned systems in the clinical arena is ca. 2 mm [ELR10, RGN13]. Such inaccuracies are still too high for the temporal-bone surgery which requires a TRE $\leq$ 0.5 mm.

An alternative to tracker-based approaches is to employ CBCT volumes to localize the instrument position with sub-voxel accuracy. After aligning the intra-operative CBCT to the pre-operative CT, it is possible to compute the instrument deviation [SMW*09]. However, the frequent acquisition of intra-operative CBCTs would result in excessive dose exposure for the patient. In order to reduce the radiation exposure, only a few X-ray images should be employed instead of 3D scans. Since single (2D) X-ray images lack 3D spatial information, it is mandatory to recover the missing depth information, e.g. by acquiring and combining at least two X-ray images from different gantry angles. Ideally these X-ray images should be acquired at positions which maximize the projection length of the first main drill axis (drill centerline), as illustrated in Figure 4.34. This assumption simplifies the localization of the surgical instrument to the detection of line-like structures. It is to be noted that only the centerline of the drill bit and the position of its tip are important to control the current instrument pose, i.e. no explicit segmentation of the projected drill bit shape is required. In summary the acquired X-ray images should be orthogonal to each other as well as to the Drilling axis.

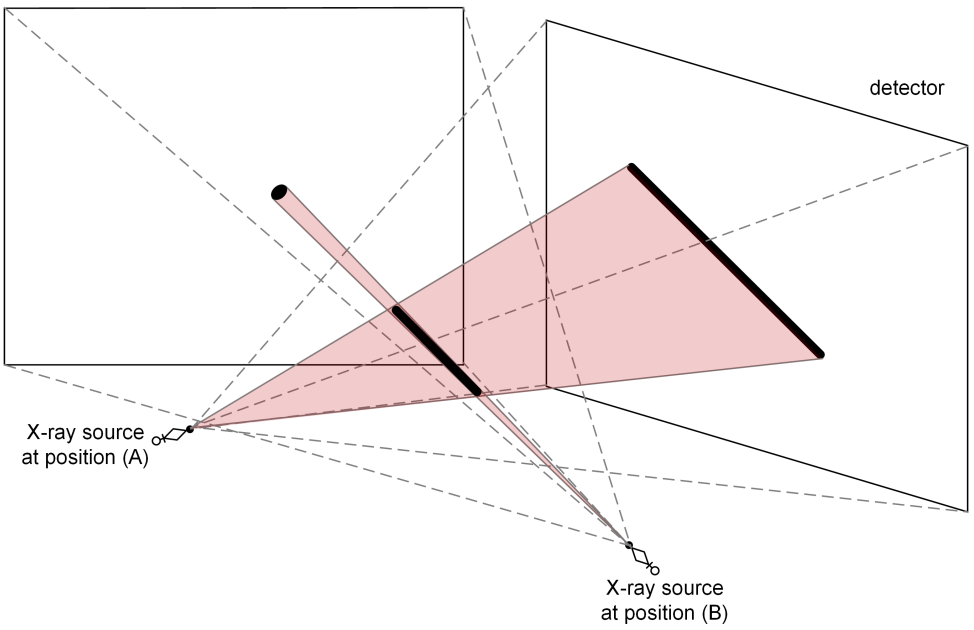


Figure 4.34.: Illustration of appropriate (A) and inappropriate (B) gantry angles for the acquisition of the intra-operative X-ray images: the projection length of the first main drill axis (drill centerline) should be maximized such as at position (B).

In order to calculate the current spatial pose of the surgical instrument, the drill bit has to be localized in both X-ray images. Several approaches have been proposed for detection of deformable surgical instruments, such as catheters [YCDS12], as well as for rigid (non-deformable) instruments, like needles used e.g. in biopsies and vertebroplasty. Due to the nature of he investigated problem, the focus is set on detection of rigid instruments appearing similar to the drill bit, such as needles.

Papalazarou et al. [PRd10] proposed a method to estimate the 3D pose of the needle during the interventional procedure based on X-ray images. This approach tries to find points belonging to the needle by following these steps:

1. preprocess the X-ray image with a bothat filter to smooth the background and enhance foreground structures
2. calculate Gaussian derivatives at multiple scales
3. find prominent maxima in scale-space, which are expected to be needle points
4. use a model fitting step to find the potential needle candidates

A drawback of this method is that many false positives can be identified as needles in case of noisy images and/or overlapping with other structures in the X-ray image.

A more robust instrument localization method has been proposed by Yanis et al. [YJ05] to localize nails (see Figure 4.35). The authors propose to first apply the Canny edge detector to the X-ray image and then use a 3D Hough transform in which the nail is modeled as a band of known width. The reported localization uncertainty is less than 1.5mm, which is insufficient for the problem under investigation. However, due to the robustness of this approach against false positives, we decided to follow it in order to find an initial estimate of the drill bit centerline in the X-ray images.

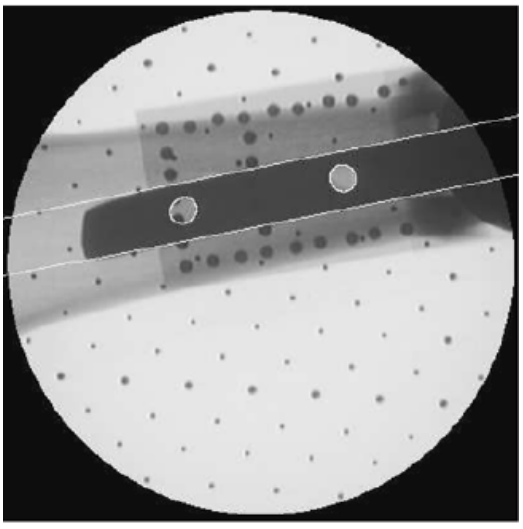
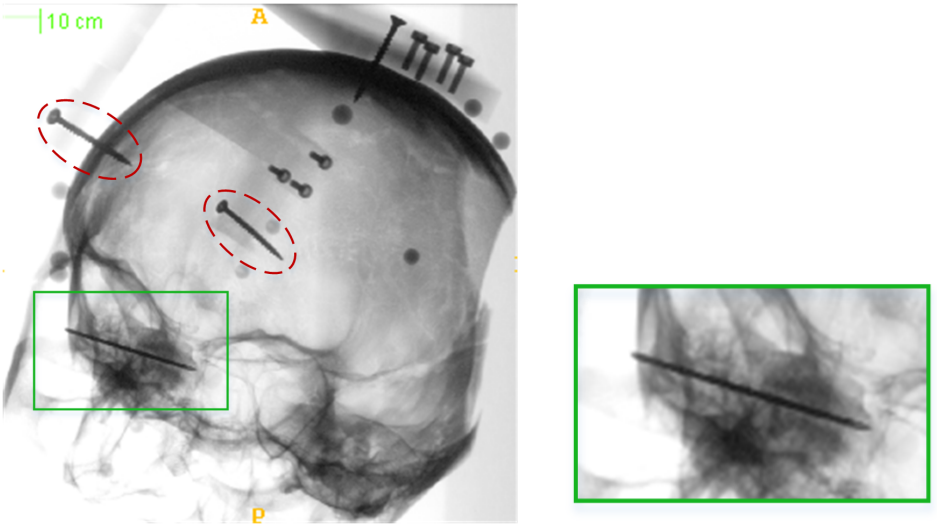


Figure 4.35.: Detected nail's longitudinal contour (white lines) in fluoroscopic X-ray image [YJ05].

4.3.1. Concept of instrument localization

The localization process starts by finding a rough estimate of the current drill bit in the intra-operatively acquired X-ray images. Usually, several structures appear similar to the drill bit, like screws shown in the red ellipses in Figure 4.36(a), which would lead to fail localizing the surgical instrument. This issue

can be partially solved by employing the planning data to define a region of interest in which the drill bit is expected to be projected. This step would reduce the number of misleading structures which look similar to the surgical instrument. Here, we assume that the drilling system is accurate enough such that the drilled channel lie close the planned path. This assumption was experimentally verified in [NHPS14]. Accordingly, a rectangle region centered at the current planned drill position is cropped as illustrated in Figure 4.36(b). For now on the cropped image region shall be used to localize the drill bit.



(a) Acquired X-ray image

(b) Cropped region which contains the drill bit

Figure 4.36.: An example showing some structures (in the red ellipses) which appear similar to the drill bit (a). A crop of a rectangular window around the planned position of the drilling instrument helps to exclude most of these misleading structures (b).

The determination of the drill bit centerline and the position of its tip (end point) consists of the following steps depicted in Figure 4.37:

1. find the boundary lines of the drill bit by employing Hough line detector (see red lines in Figure 4.37)
2. calculate the midline (blue) which lies in the middle of the two boundary lines found in step 1
3. for each pixel on the (blue) midline compute the intensity profile along the direction orthogonal to the said line and within the (red) boundary lines
4. find the intensity minimum for each intensity profile (yellow lines)
5. re-estimate the drill bit centerline by performing a linear regression based on the positions of the local minima found in fourth step
6. determine the tip of drill bit along the found centerline

These steps are described with more details below.

1) Rough localization. The drill bit might be obscured by other structures in the intra-operative X-ray images such that the drilling instrument does not appear as a continuous thick line. A simple and robust method to detect it, even when it appears as discontinuous line, is the Hough transform [DH72]. This method can be used to identify lines in an X-ray image: a straight line becomes a single point in the transformed space, which simplifies the line detection to the localization of peaks in a two dimensional parameter space. The Hough transform offers further advantages, namely

- if the drill bit is partially obscured by other structures, it is still be relatively easy to detect the boundaries since the vote of the segments (of the visible drill parts) are accumulated, which results in a peak for the corresponding boundary (line)
- criteria which are defined to select the drill bit boundary lines (see below) can be easily verified in the parameter space

However, if the Hough transform is applied to original X-ray images, several false positives would be found since the background will also contribute to the Hough accumulator. Therefore, a Canny-edge detector is employed in order to enhance discrimination of lines against the background as illustrated in Figure 4.38(b).

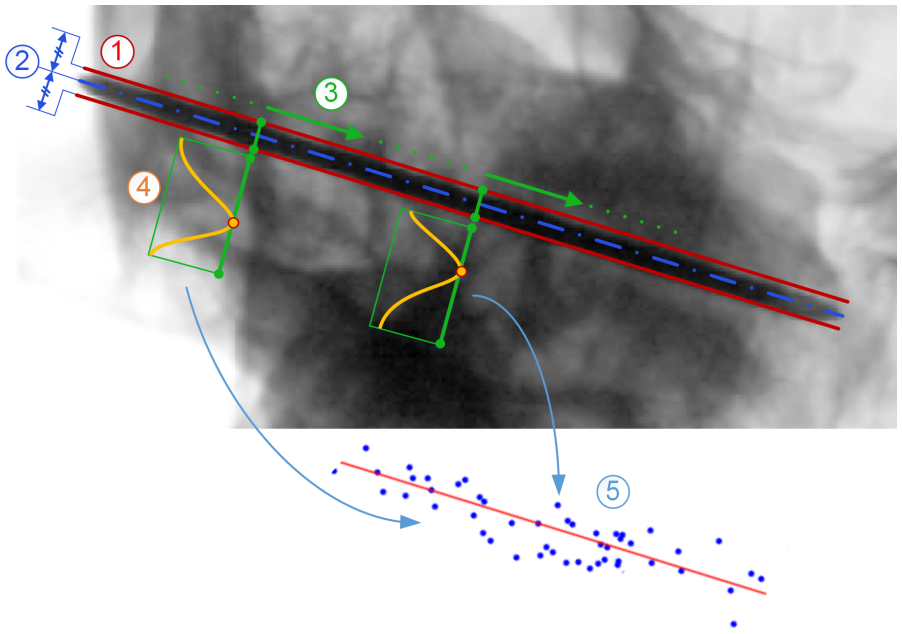


Figure 4.37.: An illustration showing the detection of the drill bit centerline: First, the boundary lines (red) are determined employing the Hough transformation. Second, the midline (blue) is calculated based on the found boundary lines. Third, an intensity profile is computed for each pixel on the midline along the directions orthogonal to the said line. Fourth, the minimum is computed for each of these profiles. Finally the found minima are used to re-estimate the drill bit centerline by applying a linear regression to the minima found the fourth step.

Ideally only two lines which correspond to the boundary of the drill bit should be found. However, in real case several potential lines are detected. Thus, it is mandatory to identify the two lines which correspond to the drill bit within all found candidates (lines). Therefore, two criteria are defined to select the two correct candidates:

- the orientation of the projection of the planned path on the X-ray image
- the diameter of the projected drill bit.

Line-pairs who satisfy these two criteria are defined as drill bit boundaries. The identification of drill bit boundary lines is performed as follows:

1. apply Canny-edge detector
2. create Hough accumulator
3. identify the peaks (each peaks correspond to a line)

4. find all peaks within the expected angle range: the planned orientation $\pm 5^{\circ}$as illustrated with the green bar in Figure 4.38(c)
5. compute the distance between each point pairs found in step 4
6. select the pair which has the expected distance as depicted in Figure 4.38(d)

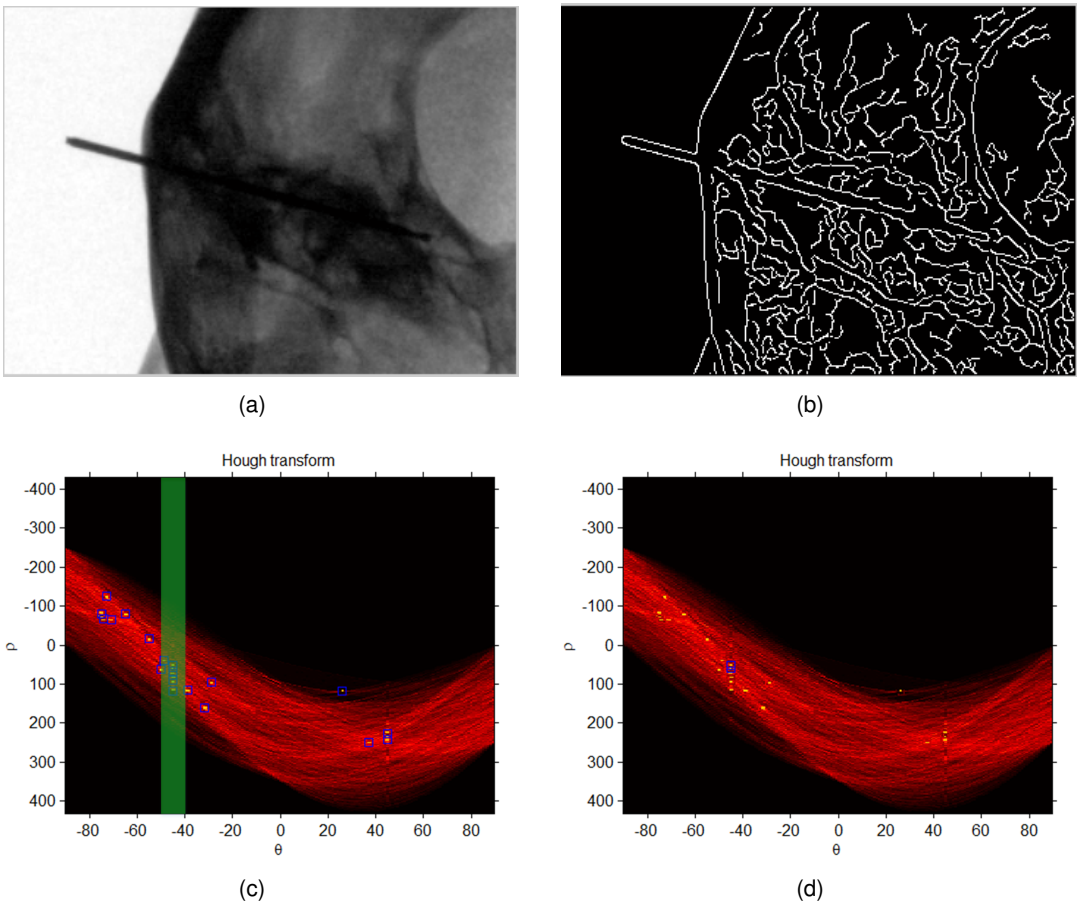


Figure 4.38.: An example showing how the boundary lines of the drill bit are identified. First the input image (a) is processed with Canny-edge detector (b), then the Hough transform is applied to it and the peaks are identified (c). Finally, the boundary lines are selected based on the expected orientation and diameter of the projection of the drill bit on the X-ray image (d).

Only one line-pair should be found, however this would not be the case when e.g. the drill bit is completely obscured (very low X-ray exposure) or another structure similar to the drill bit is projected on the planned path. In such cases the surgeon is asked to select the correct candidate or acquire another X-ray image.

2) Initial estimation of the drill bit centerline. The result of the previous rough estimation are two lines (boundaries) limiting the drill bit. An initial estimate of the drill bit centerline is defined as the line which is equidistant from the boundary lines. This estimate is sufficient but not accurate enough to be used for the navigation, thus a fine-tuning has to be done in the next steps.

3-5) Improve the estimate of the drill bit centerline. If the drill bit is not partially obscured by other structures, the X-ray attenuation is maximal along the projection ray crossing the drill bit centerline. Con-

sequently, the recorded projection values are minimal at the pixels which record the said X-rays. In order to find these pixels, we iterate over all pixels perpendicular to the initial centerline estimate. For each of these pixels, the minimum along the segment orthogonal to the initial centerline estimate and lying between the boundary lines is determined. A linear regression is then performed based on the positions of the found local minima in order to re-estimate the drill bit centerline.

A drawback of this approach is that outliers would falsify the result. This is the case if the drill bit is obscured by other metallic implants on the X-ray images such that the local minima cannot be detected or do not correspond to the drill bit centerline. In order to overcome this problem, for each segment S_D the neighbor pixels lying on this segment and are outside the boundary lines, i.e. the pixels on segments S_1 and S_2 in Figure 4.39, are considered. If the mean value of the pixels on S_1 as well as the mean value of the pixels on S_2 are both at least twice larger than the mean value of the pixels on S_D, then it is assumed that the drill bit is not obscured by other metallic objects. Only in this case the position of the minimum found for the said segment S_D is considered in the linear regression.

It is possible to perform the steps 3–5 iteratively, however in our experience one iteration is sufficient for an accurate estimation of the drill bit centerline.

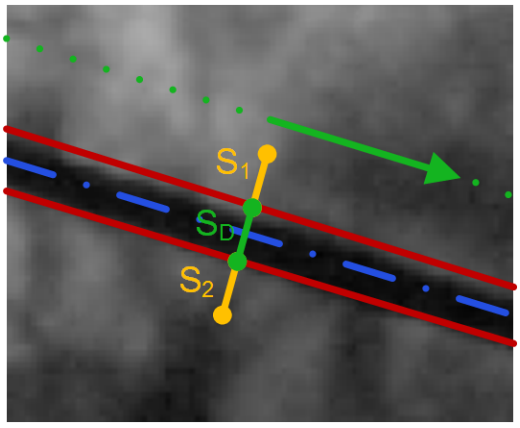


Figure 4.39.: Illustration of the method employed to detect if the drilling instrument is obscured by other metallic objects in the X-ray image: the intensity within the drill boundaries (red lines) is compared to the intensity in the surrounding regions (yellow lines).

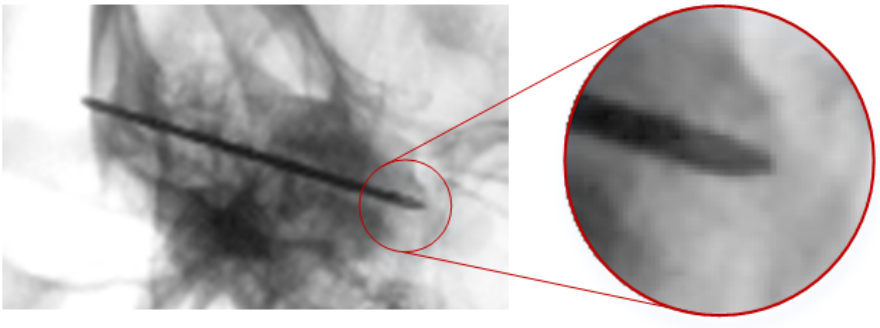


Figure 4.40.: The appearance of the drill bit tip on an X-ray image.

6) Determine the tip of drill bit. At the tip of drilling instrument the intensity should change from dark to bright with a very steep slope. Thus, the intensity profile along the centerline is employed to determine the end of drilled tunnel by looking for the steepest descent close to the expected drill bit end gained from the planning data.

The deepest point of the drilled tunnel does not always lie on the drill bit centerline as shown in Figure 4.40. In order to not miss the correct end, the detection of the tip is performed as follows:

1. iterate over all lines which are parallel to the drill bit centerline and lie between the drill boundaries
2. find the deepest drill point for each line
3. identify the global deepest point which is expected to correspond to tip of the drilling instrument

Spatial instrument pose

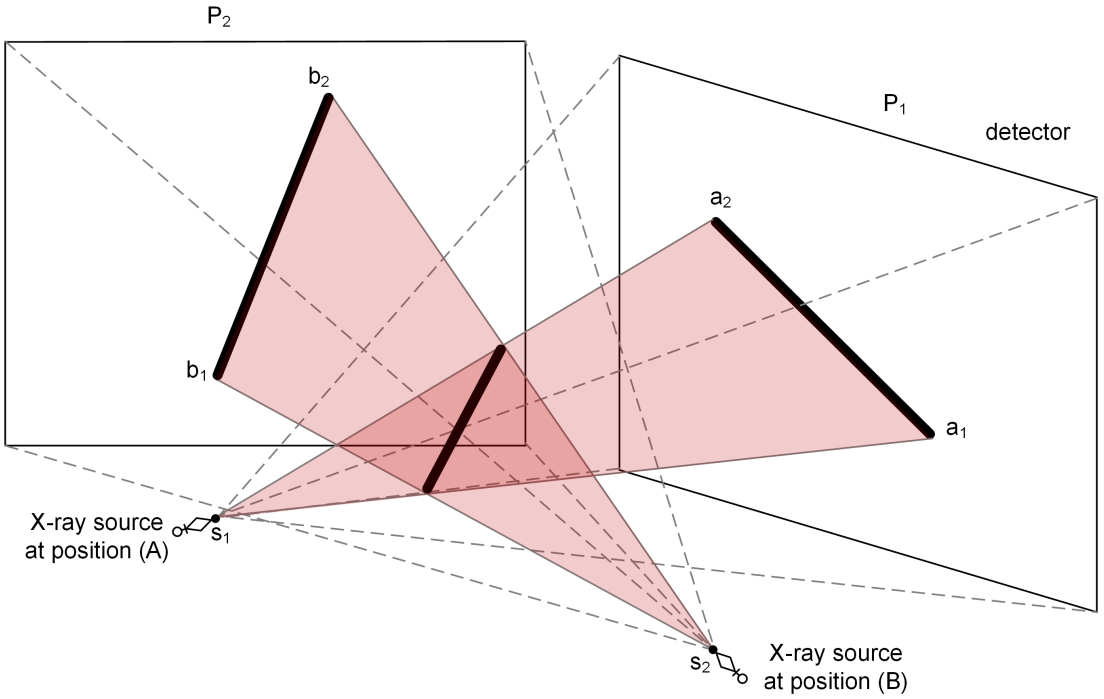


Figure 4.41.: Illustration how the spatial drill bit pose is determined: the drill bit segments on the X-ray images $\overline{a_1 a_2}$ and $\overline{b_1 b_2}$ are back-projected, which results in two triangles $\Delta(P_{focal1}, a_1, a_2)_{KS_{CT}}$ and $\Delta(P_{focal2}, b_1, b_2)_{KS_{CT}}$. The drill bit pose is defined as the intersection of this triangles. The selected points are a_2 and b_2 correspond to the tip of the drilling instrument.

At this level the projection matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ are known for both X-ray images P_1 and P_2. Furthermore, the drill bit centerline as well as its tip are localized in each X-ray image. In order to determine the (3D) pose of the drilling instrument in the coordinate system KS_{CT}, the following steps are performed (see Figure 4.41):

1. select the projection of the drill bit on each X-ray image P_1 and P_2, i.e. $\overline{a_1 a_2}$ and $\overline{b_1 b_2}$, respectively.
2. compute the position of the X-ray source relative to the CT volume at positions (A) and (B), i.e. the coordinates $(P_{focal1})_{KS_{CT}}$ and $(P_{focal2})_{KS_{CT}}$. This step can be performed with the method proposed in [NBHN*98]: Let define the 3×3 matrix $\mathbf{H}_{1a}$ and the 1×3 vector $\mathbf{H}_{1b}$ such that $\mathbf{H}_1 = [\mathbf{H}_{1a}\mathbf{H}_{1b}]$. The

position of the X-ray source at position (A) can be computed as

$$(P_{focal_1})_{KS_{CT}} = -\mathbf{H}_{1a}^{-1} \cdot \mathbf{H}_{1b} \tag{4.12}$$

Similarly, $(P_{focal_2})_{KS_{CT}}$ is calculated as

$$(P_{focal_2})_{KS_{CT}} = -\mathbf{H}_{2a}^{-1} \cdot \mathbf{H}_{2b} \tag{4.13}$$

3. compute the coordinates of the points a_1, a_2, b_1 and b_2 in KS_{CT}, i.e. $(a_1)_{KS_{CT}}$, $(a_2)_{KS_{CT}}$, $(b_1)_{KS_{CT}}$ and $(b_2)_{KS_{CT}}$. This is achieved by multiplying each of these points with the pseudo inverse of the corresponding projection matrix.
4. calculate the intersection of the triangles $\Delta(P_{focal1}, a_1, a_2)_{KS_{CT}}$ and $\Delta(P_{focal2}, b_1, b_2)_{KS_{CT}}$. The intersection segment is expected to correspond to the spatial pose of the drill bit centerline.

4.3.2. Evaluation of instrument localization accuracy

Two sets of experiments have been conducted to evaluate the accuracy of the instrument tracking algorithm, i.e. based on simulated and real data.

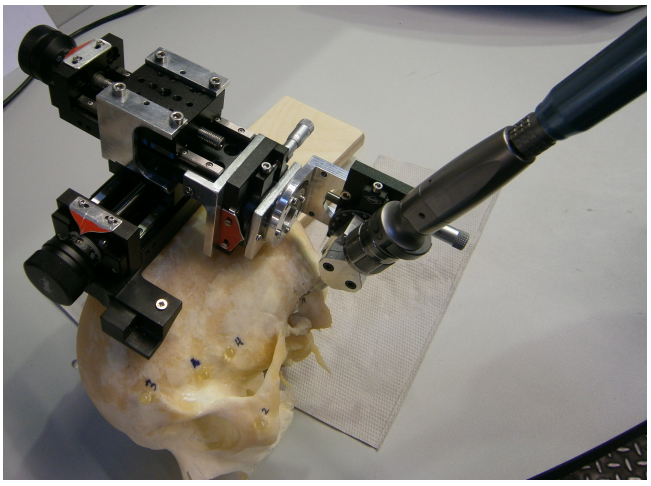


Figure 4.42.: A cadaver skull employed to evaluate the accuracy of the instrument tracking.

Simulated data. In order to check the localization accuracy, X-ray images on which a drill bit is projected as well as ground truth for the drill bit position are required, which can be easily achieved with simulated data. We have proceeded as following to create realistic data sets:

- drill a tunnel in a cadaver skull
- scan the skull (without removing the drill bit) with an industrial CBCT scanner
- segment the drill bit in the CT volume and compute its start and end points
- create DRRs form the CT volume based on geometric projections parameters similar to the C-arm CBCT system under investigation
- calculated the projection position of the drill bit on each DRR

The DRRs were provided as input for the instrument tracking. After localizing the drill instrument with the proposed algorithm, the deviation from the (know) reference position was calculated. Two error types have been evaluated:

- ϵ_d: the Euclidean distance between the estimated endpoint (on the centerline) and the reference position of the drill bit tip
- ϵ_α: the inner angle between the estimated centerline and the reference centerline

In ideal case both errors are expected to be zero. It is to be mentioned that over 100 DRRs were created while respecting the assumptions listed at the beginning of this section. The achieved mean ϵ_d was 0.092 mm and the mean ϵ_α was 0.181° which is acceptable for the target application. The results are summarized in 4.43.

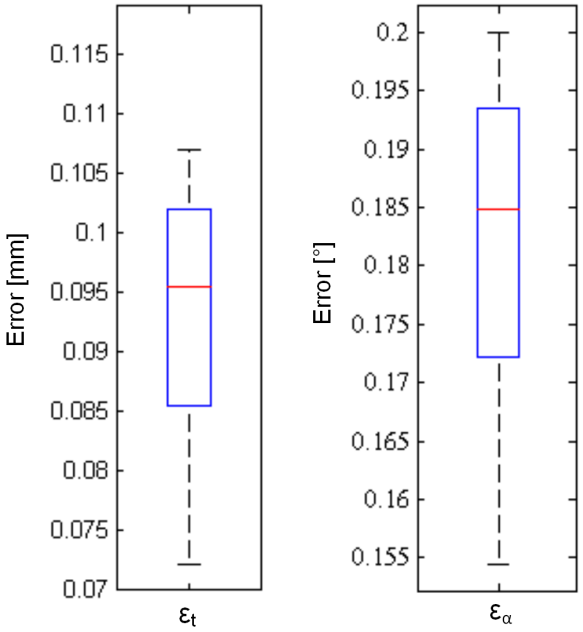
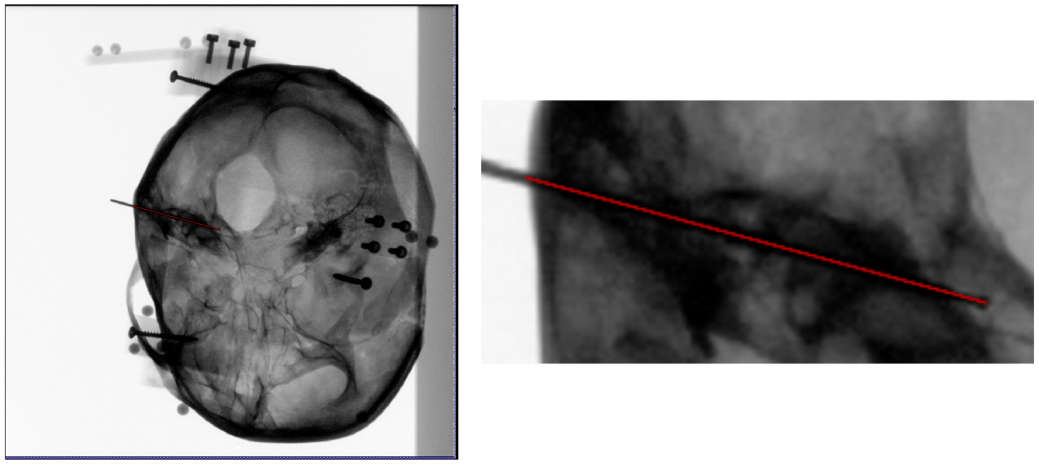


Figure 4.43.: The localization error based on simulated data: ϵ_d is the Euclidean distance between the estimated endpoint and the correct position, $\epsilon\alpha$ is the inner angle between the estimated and the reference centerlines.

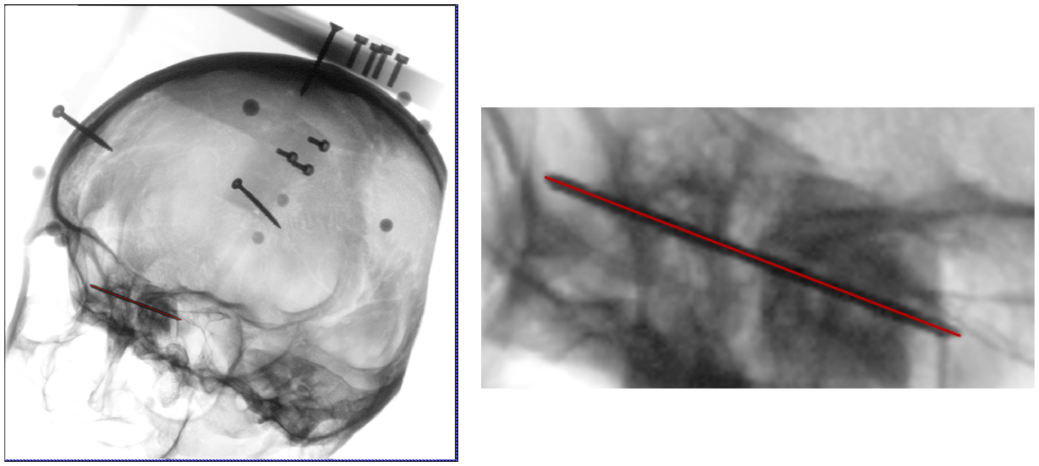
Real X-ray images. The proposed tracking method was tested on 24 X-ray images. In this case even manual segmentation would not be accurate enough for quantitative evaluation. Due to the missing ground-truth, the evaluation was limited to visual inspection. In all processed images, the localization accuracy was good. Some examples are shown in Figure 4.44.

4.4. Navigation

The introduction of C-arm systems allowed the further concomitant growth in the field of computer-aided interventions. It has been broadly studied and applied to different applications such as orthopedic and spinal surgery [LD04, YCT*12, LHWG12, CLT*15]. Some approaches have been also investigated for head and neck surgery [LD04, YCT*12, LOA*14]. For instance, Liu et al. proposed an image guidance system for video-based minimally-invasive robotic surgery. The system integrated a high-end robotic C-arm system



(a)



(b)

Figure 4.44.: Two examples illustrating the accuracy by localizing the drill bit centerline and tip in X-ray images using the proposed tracking method.

with the da Vinci platform [LOA*14]. The authors reported a mean target registration error $TRE_{mean} \leq 2.0$ mm using a Siemens Artis zeego (2012, Siemens Inc.) equipped with a 30×40 cm flat-panel detector.

Unlike orthopedic and spinal surgery where intra-operative surgical guidance is common [LD04, YCT*12, LHWG12, CLT*15], intra-operative image guidance for head and neck surgery has had a relatively limited experience [LD04, YCT*12, LOA*14, LBN*14], and even more limited work has been proposed for minimally invasive image-guided cochlear implantation [BSG*12, LBN*14]. Labadie et al. [LBN*14] demonstrated percutaneous single-port access to the cochlea in vitro and in vivo using preoperative computed tomography, bone-implanted fiducial markers, and a customized microstereotactic frame. The control of the drilling course is performed by acquiring intra-operative CT volumes. Due to the high dose exposure, the image-guided drilling control was performed only twice: in the begin of the process and after the lateral drilling. The drawback of this approach is that a frequent control of the drilling process is not possible due to the high dose exposure.

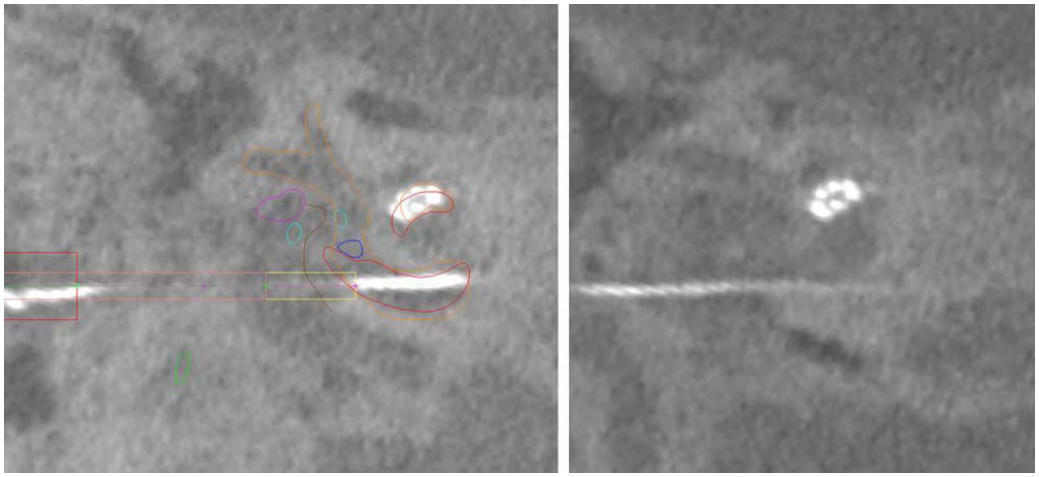


Figure 4.45.: Postoperative screenshot showing the tip fold-over [LBN*14].

Bell et al. [BSG*12] presented an image-guided robot for drilling of one access tunnel for the insertion of an electrode in temporal bone cadaver specimens. The used robot was mounted to the operation table, whereas the specimen was immobilized with head clamps. Reference marker screws were attached to the skull and used for registration: the positions of these markers were measured by the robot in order to align it to the specimen. A single CBCT was acquired in a proxi-operative stage for the planning as well as for navigation. This would be sufficient under the assumption that the relative pose of all system parts remain unchangeable. The authors reported a mean accuracy at the target point of 0.56 $\pm$41 mm with an angular misalignment of 0.88° $\pm$0.41°. The main drawback of this method is the non-intraoperative control of the drill process. The navigation is obviously based on the assumption that the accuracy of the robot system is high enough to achieve the drilling process with high precision. Moreover, the necessity for an additional device (robot) imposes constraints to the clinical application.

Instrument positioning error. At this level the instrument position relative to the planned path is known. In order to calculate the positioning error, these two comparison are performed:

- the Euclidean distance between the current and the planned position of the drill bit tip
- the inner angle between the current and the planned centerline of the drilling instrument

In ideal case both errors should be equal to zero.

Visualization. The deviation of the current drill pose from the planned drilling path is visualized by highlighting the drill bit centerline and projecting the planned path on the X-ray images. As illustrated in Figure 4.46, the planned path (green line) does not cover the drill bit (red line), which indicates a relatively high positioning error.

4.5. Summary and conclusions

A hybrid registration approach has been presented to align intra-operative X-ray images, acquired with a C-arm CBCT system, relative to pre-operative CT volume. This task is crucial for the image-guided navigation technique which was proposed to control the drilling process. A marker-based approach has been introduced to find an initial estimate for the C-arm pose. This estimate is refined by employing an

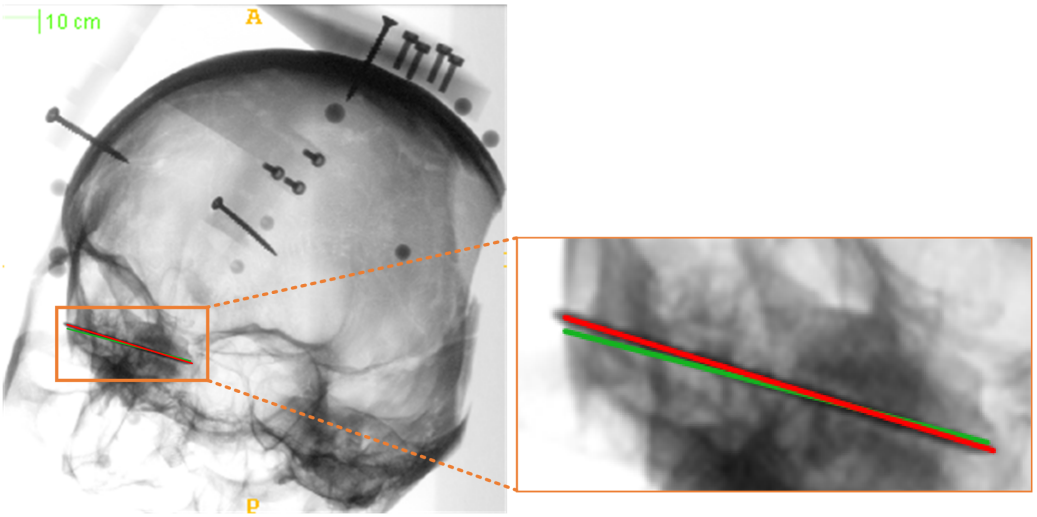


Figure 4.46.: Visualization of the current drill bit position and the planned drilling path on an X-ray image.

intensity-based registration. No specific configuration was assumed for the used fiducial markers, which offers more flexibility in placing the fiducials. The proposed registration is computationally intensive, mainly due the creation of DRRs. This part was therefore performed on GPU in order to reduce the processing time. The efficiency and accuracy of the proposed method have been demonstrated.

The variability in fiducial appearance, the presence of noise and superimposed structures make it difficult to accurately localize fiducial centers in X-ray images. An accurate, robust, and fast method has been proposed to achieve this task. The introduced particle path segmentation method requires merely a rough initial center candidate located somewhere within the marker contour. It can improve the fiducial localization accuracy by considering prior knowledge about the marker shape and estimating the reliability of the input information. Therefore, the target registration error could be reduced by an order of magnitude compared to alternative popular approaches. The proposed particle path segmentation has been designed for circular shapes, but it can be extended to segment other more or less circular objects (like cells or nuclei) as well as convex objects.

After tracking the drill bit in both X-ray images, its spatial pose is computed by employing the calculated C-arm poses for both images. Then, a comparison with the planned path is performed to calculate the drill bit positioning error. Experiments based on simulated and real data have demonstrated the accuracy of the proposed methods.

5. Conclusions and future work

This thesis presented new methods for accurate 3D CT reconstruction and image-guided navigation. This chapter provides a summary of the contributions in Section 5.1, states open problems and suggests new ideas for future research in Section 5.2.

5.1. Summary

The overall aim of the project MUKNO is the investigation of the feasibility of minimally traumatic methods while considering all possible error sources and their quantification in order to predict whether the required accuracy can be achieved for a specific patient [NHSB*14]. In this context it is to be considered that the whole treatment pipeline is built on the acquired medical image data. Therefore, it is essential to acquire the images with high accuracy in order to allow the creation of precise patient specific anatomical models. For high precision surgery at the lateral skull base, accurate navigation is required to track the current instrument position and eventually correct its location. These two major fields, i.e. 3D CT reconstruction and image-guided navigation, have been studied in this thesis.

5.1.1. Super-resolution reconstruction (SRR)

3D CT images acquired with spiral CT scanners, referred here as low-resolution images, have typically anisotropic resolution. A new SRR approach has been proposed to recover a volume with isotropic voxels from only two low-resolution images. The concept of confidence maps has been introduced in order to incorporate prior knowledge about the acquisition process, such as noise level. It aims to estimate the confidence of the input images and then recover the unknown high-resolution volume by maximizing the data similarity to the high-confident information. The proposed method allows the reconstruction of HR volumes from already acquired 3D CT images. This offers the possibility to fuse data sets acquired with different scanners and/or over different times. The method is fully-automatic: it does not require any user interaction for the alignment of the input LR images and the reconstruction of the HR volume. Furthermore, no additional hardware is required, which allows easy inclusion of the proposed technique in clinical suites.

The evaluation of the contributed methods has shown their capability to reconstruct volumes with isotropic resolution while recovering high-frequency details and fine anatomical structures. It is to be noted that the image resolution should not be mistaken for imaging accuracy. This aspect was the focus of the segmentation based evaluation: it was shown that the segmentation uncertainty was reduced to achieve a sub-voxel accuracy in all directions if the anisotropy factor does not exceed the factor five, which is a standard in modern spiral CT scanners.

5.1.2. CBCT reconstruction

This thesis contributed new algorithms for CBCT reconstruction. Two major problems have been addressed: mechanical inaccuracies (jittering) and metal-artifacts.

Mechanical inaccuracies. CBCT volumes are reconstructed based on X-ray images gathered during the rotation of the C-arm around the object under investigation. It is mandatory to provide accurate geometrical projection parameters for each acquired X-ray image, otherwise significant CT artifacts may be induced in the reconstructed volume. Usually, a pre-calibration (offline calibration) using an X-ray phantom is preformed under the assumption that the C-arm motion is reproducible. Thereby, stochastic misalignments due to the open design of mobile C-arm CBCT systems are not considered. A novel online calibration algorithm has been introduced to compensate stochastic mechanical inaccuracies. The method can be applied intra-operatively since it does not require additional hardware and assumes only a pre-calibration (offline calibration), i.e. the remaining misalignment error is rather small (around 1 mm). After performing a sensitivity study for the scanner under investigation, we have shown that the calibration process can be simplified by reducing the number of the estimated geometric projection parameters without causing significant degradation in the reconstruction quality. This can significantly reduce computation times, which makes the method efficient for intra-operative applications. Furthermore, the influence of metallic artifacts has been suppressed in the process. Using the proposed on-the-fly calibration and enhanced similarity measure, the accuracy and quality of reconstructed CBCTs could be improved. Experiments based on simulated and real data have demonstrated that the method can achieve a sub-pixel accuracy at least for the investigated C-arm device.

Metal artifact reduction. The presence of high-density objects, like metal parts, induces streak-like artifacts, which can obscure relevant anatomy. A novel non-linear back-projection kernel for iterative CBCT reconstruction has been proposed to reduce such artifacts and enhance the quality of reconstructed 3D images. The reconstruction algorithm has been evaluated with several CBCT systems. It could achieve better reconstruction results than the standard techniques. Excessive metal artifacts were even completely eliminated and details lost in the corrupted regions were recovered. The main reason for this enhanced reconstruction quality is the consideration of the variance of the projection information simultaneously over all projection images. Good results were computed within 2 to 3 iterations, in contrast to other methods which may require significantly more iterations. Such enhancements should advance clinical applications of the C-arm CBCT, like diagnosis, navigation, and visualization.

5.1.3. Image-guided navigation

An image-guided navigation has been proposed. For this task a C-arm CBCT system, which is standard element in modern operation rooms, is employed in combination with some fiducial markers attached to the patient's skull. The system does not require any additional sensors, such as electromagnetic sensor, and is therefore not manufacturer-dependent.

The contribution in this context was focused on the estimation of the marker center points since the accuracy of this step is decisive for the achievable overall navigation precision. Furthermore, performing this

task automatically would speed up the registration process and reduce the human error. However, a sub-pixel accurate localization of the said centers is at present a challenging task, especially when fiducials overlap dense anatomical structures on the radiographs. A new method, called particle path segmentation (PPS), has been proposed for automated and accurate localization of marker centers in X-ray images, even in presence of structure overlapping. It requires merely a rough initial center candidate located somewhere within the marker contour.

Several experiments have confirmed the high accuracy and robustness of the proposed algorithm achieving a localization error –mean (std)– equal to 0.059 mm (0.062), i.e. an order of magnitude compared to alternative popular approaches. A such improvement could be achieved by considering prior knowledge about the marker shape and estimating the reliability of input information for segmentation. The proposed algorithm has been designed for circular shapes, but can be extended to segment any convex object.

In addition, approaches have been proposed to localize the drilling instrument in X-ray images in order to determine its spatial position. This serves to calculate the deviation of the current drill bit pose from the planned path. In such a way the drill process can be controlled and the surgeon can be supported by visualizing the current positioning error on the X-ray images.

5.2. Future Work

It is often not possible to perform a full scan rotation (around the object under investigation) due to mechanical limitations. A typical scan trajectory is a circular arc with a definite angle range between 135° and 165°. In order to acquire data from the missing directions, it is possible to perform two scans while changing the pose of the patient relative to the scanner between both scans, either by rotating the patient or the scanner itself. Thereby the total radiation exposure can be kept in the same range by acquiring only the half of X-ray projections in each scans. In this case, both data sets (X-ray images) have to be registered to each others. This can be achieved by employing at least three fiducials to perform point-based registration. An important question to be answered is how to incorporate redundant data in the reconstruction algorithm to increase the signal to noise ratio.

The proposed CBCT reconstruction algorithm is limited to rigid anatomies. It would be interesting to extend it for deformable objects, such as cardiac CT where motion artifacts are expected since the heart is moving during the acquisition of the X-ray images.

The presented image-guided navigation should be applied intra-operatively to correct the instrument position and not only to monitor the deviation. However, such a scheme is not possible with the current drill hardware. In addition, the planned trajectories are limited to linear tunnels. Initial Experiments conducted on cadaver skulls have shown that under this constrain not enough manipulation room is possible. The drill of non-linear paths would offers more freedom and security in defining the tunnels while maximizing the distance to the nearest critical structure. Tools allowing the drill of non-linear channels have therefore to be first developed and the navigation system has to be accordingly adapted.

For a continuous controlling of the drilling process, a hybrid system would be required since the proposed image-guided navigation should not be employed continuously, otherwise it would result in high radiation exposure for the patient. In addition, it is not possible to accurately estimate the rotation around the central

axis (roll) with the C-arm device. Thus, we propose to combine the proposed C-arm guided navigation with an electromagnetic tracking system in order to increase sampling-rate.

A. Publications and talks

The thesis is partially based on the following publications:

1. EL HAKIMI W., SAKAS G.: On-the-fly geometrical calibration fine-tuning of a mobile C-arm CBCT system. *Journal Technisches Messen: Medical Metrology – An interdisciplinary perspective*, (Accepted) (2015)
2. EL HAKIMI W., BEUTEL J., SAKAS G.: Particle Path Segmentation: a Fast, Accurate, and Robust Method for Localization of Spherical Markers in Cone-beam CT Projections. *DGBMT: Biomedical engineering 59 Suppl 1* (2014), s405–408
3. EL HAKIMI W., SAKAS G.: Variance-based iterative image reconstruction from few views in limited-angle C-arm computed tomography. In *SPIE Medical Imaging* (2014), Whiting B. R., Hoeschen C., (Eds.), SPIE Proceedings, SPIE, p. 90332Q
4. STEGER T., EL-HAKIMI W., WESARG S.: Tracking von Instrumenten auf fluoroskopischen Aufnahmen für die navigierte Bronchoskopie. In *Bildverarbeitung für die Medizin 2014*, Deserno T. M., Handels H., Meinzer H.-P., Tolxdorff T., (Eds.), Informatik aktuell. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 228–233
5. IGOR STENIN, STEFAN HANSEN, MARIA NAU-HERMES, WISSAM EL HAKIMI, MEIKE BECKER, JUDITH BEVERMANN, THOMAS KLENZNER, JOERG SCHIPPER: Evaluation von minimal invasiven multiport Zugängen der Otobasis am humanen Schädelpräparat. In *Hubertus Feussner (Ed.) 2014 – 13. Jahrestagung der Deutschen Gesellschaft*. pp. 127–130
6. NAU-HERMES M., SCHMITT R., BECKER M., EL-HAKIMI W., HANSEN S., KLENZNER T., SCHIPPER J.: Quality assurance of multiport image-guided minimally invasive surgery at the lateral skull base. *BioMed research international 2014* (2014), 904803
7. EL HAKIMI W., WESARG S.: Accurate super-resolution reconstruction for CT and MR images. In *26th IEEE International Symposium on Computer-Based Medical Systems (CBMS)* (2013), pp. 445–448
8. KEYHANI M. H., HAKIMI W. E., WESARG S.: Anisotropy correction of medical image data employing patch similarity. In *2013 IEEE 26th International Symposium on Computer-Based Medical Systems (CBMS)* (2013), pp. 385–388
9. EL HAKIMI W., WESARG S.: Confidence map based super-resolution reconstruction. In *SPIE Medical Imaging* (2012), Haynor D. R., Ourselin S., (Eds.), SPIE Proceedings, SPIE, p. 831440
10. WESARG S., BECKER M., EL HAKIMI W.: Planning of High Precision Surgery at the Lateral Skull Base. In *14th Korea-Germany Joint Workshop on Advanced Medical Image Processing 2011. Heidelberg, Germany*. 2011, pp. 72–78

B. Supervising activities

The following list summarizes the student bachelor, diploma and master thesis supervised by the author. The results of these works were partially used as an input into the thesis.

B.1. Diploma and master thesis

1. FRIESS B.: Automatische Bestimmung der optimalen Parameter für intraoperative Bildaufnahmen und Instrumentenverifikation, 2014
2. BEUTEL J.: Sub-Pixel Accurate Fiducial Marker Detection and Pose Estimation from CBCT-Data, 2013
3. HERBERT S.: Super-Resolution Reconstruction for Medical Image Enhancement using Uncertainty Models, 2011

B.2. Bachelor thesis

1. MEHLTRETTER M.: Optimierung einer 2D/3D Registrierungsmethodik für bildgestützte Eingriffe, 2014
2. DROST Y.: Bildrekonstruktion aus Kegelstrahlprojektionen für zwei Scan-Trajektorien, 2014

C. Curriculum Vitae

Personal Data

Name	Wissam El Hakimi
Birth date & place	04.03.1985 in Le Kef, Tunisia
Nationality	Tunisian, German

Education

2010	Master of Science in Automation Technology, University of Stuttgart, Germany
2005	DSH (language proficiency), University of Heidelberg, Germany
2004	Technological Baccalaureate, Lycée Pilote du Kef, Tunisia

Work Experience

2011 – 2016	Researcher, Interactive Graphics Systems Group, Technische Universität Darmstadt, Germany, Focus: Medical Computing
2010	Student software developer at Bosch Rexroth AG, Stuttgart
2007 – 2009	Student research assistant at IFF, University of Stuttgart

Bibliography

[AFCS11] ARIAS P., FACCIOLO G., CASELLES V., SAPIRO G.: A Variational Framework for Exemplar-Based Image Inpainting. *International Journal of Computer Vision 93*, 3 (2011), 319–347. 30

[AFP91] ALGAZI V. R., FORD G. E., POTHARLANKA R.: Directional interpolation of images based on visual properties and rank order filtering. In *[Proceedings] ICASSP 91: 1991 International Conference on Acoustics, Speech, and Signal Processing* (1991), pp. 3005–3008 vol.4. 31, 117

[ALD*10] AYAD M. S., LEE J., DEGUET A., BURDETTE E. C., PRINCE J. L.: C-arm Pose Estimation Using a Set of Coplanar Ellipses in Correspondence. *Proceedings / IEEE International Symposium on Biomedical Imaging: from nano to macro. IEEE International Symposium on Biomedical Imaging 2010* (2010), 1401–1404. 134, 135

[BBB*07] BRANDENBERGER D., BIRKFELLNER W., BAUMANN B., MESSMER P., HUEGLI R. W., REGAZZONI P., JACOB A. L.: Positioning accuracy in a registration-free CT-based navigation system. *Physics in medicine and biology 52*, 23 (2007), 7073–7086. 130

[BDLT10] BENSON T. M., DE MAN,B K B, LIN FU, THIBAULT J.-B.: Block-based iterative coordinate descent. In *2010 IEEE Nuclear Science Symposium and Medical Imaging Conference (2010 NSS/MIC)* (2010), pp. 2856–2859. 68, 75

[BDS09] BODENSTEINER C., DAROLTI C., SCHWEIKARD A.: Super-Resolution für mobile C-Bogen-Systeme. In *Bildverarbeitung für die Medizin 2009*, Brauer W., Meinzer H.-P., Deserno T. M., Handels H., Tolxdorff T., (Eds.), Informatik aktuell. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 152–156. 18

[Bec14] BECKER M.: *Patientenspezifische Planung für die Multi-Port Otobasischirurgie*. PhD thesis, Darmstadt, TU, Diss., 2014, 2014. v, 124, 125

[Beu13] BEUTEL J.: Sub-Pixel Accurate Fiducial Marker Detection and Pose Estimation from CBCT-Data, 2013. 175

[BH07] BRENNER D. J., HALL E. J.: Computed tomography–an increasing source of radiation exposure. *The New England journal of medicine 357*, 22 (2007), 2277–2284. 77

[BJF07] BURKHARDT D., JAIN A., FICHTINGER G.: A cheap and easy method for 3D C-arm reconstruction using elliptic curves. In *Medical Imaging* (2007), Cleary K. R., Miga M. I., (Eds.), SPIE Proceedings, SPIE, pp. 65090B–65090B–10. 134

[BK02] BAKER S., KANADE T.: Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence 24*, 9 (2002), 1167–1183. 8

[BKK12] BEISTER M., KOLDITZ D., KALENDER W. A.: Iterative reconstruction methods in X-ray CT. *Physica medica : PM : an international journal devoted to the applications of physics to medicine and biology : official journal of the Italian Association of Biomedical Physics (AIFB) 28*, 2 (2012), 94–108. 67

[BLD06] BERNHARDT P., LENDL M., DEINZER F.: New technologies to reduce pediatric radiation doses. *Pediatric Radiology 36*, Suppl 2 (2006), 212–215. 18

[BM92] BESL P. J., MCKAY H. D.: A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence 14*, 2 (1992), 239–256. 24

[BO07] BUZUG T., OEHLER M.: Statistical image reconstruction for inconsistent CT projection data. *Methods of information in medicine 46*, 3 (2007), 261–269. 117

[BR67] BRACEWELL R. N., RIDDLE A. C.: Inversion of Fan-Beam Scans in Radio Astronomy. *The Astrophysical Journal 150* (1967), 427. 66

[BS04] BRUNO DE MAN, SAMIT BASU: Distance-driven projection and backprojection in three dimensions. *Physics in medicine and biology 49*, 11 (2004), 2463. 69

[BS06] BAL M., SPIES L.: Metal artifact reduction in CT using tissue-class modeling and adaptive prefiltering. *Medical physics 33*, 8 (2006), 2852–2859. 78, 79, 116

[BSCB00] BERTALMIO M., SAPIRO G., CASELLES V., BALLESTER C.: Image inpainting. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (2000), pp. 417–424. 117

[BSF*09] BIRKFELLNER W., STOCK M., FIGL M., GENDRIN C., HUMMEL J., DONG S., KETTENBACH J., GEORG D., BERGMANN H.: Stochastic rank correlation: A robust merit function for 2D/3D registration of image data obtained at different energies. *Medical Physics 36*, 8 (2009), 3420. 131

[BSG*12] BELL B., STIEGER C., GERBER N., ARNOLD A., NAUER C., HAMACHER V., KOMPIS M., NOLTE L., CAVERSACCIO M., WEBER S.: A self-developed and constructed robot for minimally invasive cochlear implantation. *Acta oto-laryngologica 132*, 4 (2012), 355–360. 166, 167

[CBNC09] COUSTY J., BERTRAND G., NAJMAN L., COUPRIE M.: Watershed cuts: minimum spanning forests and the drop of water principle. *IEEE transactions on pattern analysis and machine intelligence 31*, 8 (2009), 1362–1374. 140, 154

[CBW*07] CHERYAUKA A., BARRETT J., WANG Z., LITVIN A., HAMADEH A., BEAUDET D.: 3-D geometry calibration and markerless electromagnetic tracking with a mobile C-arm. *Proc. SPIE 6509* (2007). 82

[CDS08] COWEN A. R., DAVIES A. G., SIVANANTHAN M. U.: The design and imaging characteristics of dynamic, solid-state, flat-panel x-ray image detectors for digital fluoroscopy and fluorography. *Clinical radiology 63*, 10 (2008), 1073–1085. 55

[CEGL10] COLONNA F., EASLEY G., GUO K., LABATE D.: Radon transform inversion using the shearlet representation. *Applied and Computational Harmonic Analysis 29*, 2 (2010), 232–250. 92

[Chr08] CHRISTIAN SCHMIDGUNST: Modell zur Konditionierung von Festkörperdetektoren für die 2D/3D-Röntgenbildgebung mit mobilen C-Bogensystemen, 2008. 56

[CKK*07] CITAK M., KENDOFF D., KFURI M., PEARLE A., KRETTEK C., HÜFNER T.: Accuracy analysis of Iso-C3D versus fluoroscopy-based navigated retrograde drilling of osteochondral lesions: a pilot study. *The Journal of bone and joint surgery. British volume 89*, 3 (2007), 323–326. 123

[CLT*15] CHANG C.-J., LIN G.-L., TSE A., CHU H.-Y., TSENG C.-S.: Registration of 2D C-Arm and 3D CT Images for a C-Arm Image-Assisted Navigation System for Spinal Surgery. *Applied Bionics and Biomechanics 2015* (2015), 1–9. 165, 166

[Col77] COLSHER J. G.: Iterative three-dimensional image reconstruction from tomographic projections. *Computer Graphics and Image Processing 6*, 6 (1977), 513–537. 67

[CP10] CLEARY K., PETERS T. M.: Image-guided interventions: technology review and clinical applications. *Annual review of biomedical engineering 12* (2010), 119–142. 123

[Cri04] CRISTIAN BADEA AND RICHARD GORDON: Experiments with the nonlinear and chaotic behaviour of the multiplicative algebraic reconstruction technique (MART) algorithm for computed tomography. *Physics in medicine and biology 49*, 8 (2004), 1455. 72

[CS05] CHAN T. F., SELIM ESEDO: Aspects of total variation regularized L 1 function approximation. *SIAM J. APPL. MATH* (2005). 35

[CZK*98] COLLINS D. L., ZIJDENBOS A. P., KOLLOKIAN V., SLED J. G., KABANI N. J., HOLMES C. J., EVANS A. C.: Design and construction of a realistic digital brain phantom. *IEEE transactions on medical imaging 17*, 3 (1998), 463–468. 48

[DC94] DEFRISE M., CLACK R.: A cone-beam reconstruction algorithm using shift-variant filtering and cone-beam backprojection. *IEEE transactions on medical imaging 13*, 1 (1994), 186–195. 92

[DDDS04] DAVID P., DEMENTHON D., DURAISWAMI R., SAMET H.: SoftPOSIT: Simultaneous Pose and Correspondence Determination. *International Journal of Computer Vision 59*, 3 (2004), 259–284. 152

[DH72] DUDA R. O., HART P. E.: Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM 15*, 1 (1972), 11–15. 159

[dPAB90] DEL VALLE M., POCH M., ALONSO J., BARTROLI J.: Comparison of the Powell and simplex methods in the optimization of flow-injection systems. Simulation on modelled experimental surfaces and experimental optimizations. *Analytica Chimica Acta 241*, 1 (1990), 31–42. 86, 152

[Dro14] DROST Y.: Bildrekonstruktion aus Kegelstrahlprojektionen für zwei Scan-Trajektorien, 2014. 175

[DSC*08] DALY M. J., SIEWERDSEN J. H., CHO Y. B., JAFFRAY D. A., IRISH J. C.: Geometric calibration of a mobile C-arm for intraoperative cone-beam CT. *Medical Physics 35*, 5 (2008), 2124. 57, 82

[DW82] D C YOULA, WEBB: H.: Image restoration by the method of convex projections: Part 1 – theory. *IEEE Trans. Med. Imaging*, 1 (1982), 81–94. 11, 16

[EBS14] EL HAKIMI W., BEUTEL J., SAKAS G.: Particle Path Segmentation: a Fast, Accurate, and Robust Method for Localization of Spherical Markers in Cone-beam CT Projections. *DGBMT: Biomedical engineering 59 Suppl 1* (2014), s405–408. viii, 143

[EF97] ELAD M., FEUER A.: Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images. *Image Processing, IEEE Transactions on 6*, 12 (1997), 1646–1658. 11, 16, 17

[EF99] ERDOGAN H., FESSLER J. A.: Ordered subsets algorithms for transmission tomography. *Physics in medicine and biology 44*, 11 (1999), 2835–2851. 76

[ELR10] ELFRING R., LA FUENTE M. D., RADERMACHER K.: Assessment of optical localizer accuracy for computer aided surgery systems. *Computer aided surgery : official journal of the International Society for Computer Aided Surgery 15*, 1-3 (2010), 1–12. 157

[ES14] EL HAKIMI W., SAKAS G.: Variance-based iterative image reconstruction from few views in limited-angle C-arm computed tomography. In *SPIE Medical Imaging* (2014), Whiting B. R., Hoeschen C., (Eds.), SPIE Proceedings, SPIE, p. 90332Q. vii, 81

[ES15] EL HAKIMI W., SAKAS G.: On-the-fly geometrical calibration fine-tuning of a mobile C-arm CBCT system. *Journal Technisches Messen: Medical Metrology – An interdisciplinary perspective*, (Accepted) (2015). vii, 81

[EW12] EL HAKIMI W., WESARG S.: Confidence map based super-resolution reconstruction. In *SPIE Medical Imaging* (2012), Haynor D. R., Ourselin S., (Eds.), SPIE Proceedings, SPIE, p. 831440. vi, 8

[EW13] EL HAKIMI W., WESARG S.: Accurate super-resolution reconstruction for CT and MR images. In *26th IEEE International Symposium on Computer-Based Medical Systems (CBMS)* (2013),

pp. 445–448. vi, 8

[FDK84] FELDKAMP L. A., DAVIS L. C., KRESS J. W.: Practical cone-beam algorithm. *Journal of the Optical Society of America A 1*, 6 (1984), 612. 53, 66

[Fia01] FIAT D.: Method of enhancing an MRI signal, 2001. 17

[Fit03] FITZGIBBON A. W.: Robust registration of 2D and 3D point sets. *Image and Vision Computing 21*, 13-14 (2003), 1145–1153. 152

[FJP02] FREEMAN W. T., JONES T. R., PASZTOR E. C.: Example-based super-resolution. *IEEE Computer Graphics and Applications 22*, 2 (2002), 56–65. 17

[Fri14] FRIESS B.: Automatische Bestimmung der optimalen Parameter für intraoperative Bildaufnahmen und Instrumentenverifikation, 2014. 125, 175

[FST*13] FU J., SCHLEEDE S., TAN R., CHEN L., BECH M., ACHTERHOLD K., GIFFORD M., LOEWEN R., RUTH R., PFEIFFER F.: An algebraic iterative reconstruction technique for differential X-ray phase-contrast computed tomography. *Zeitschrift für medizinische Physik 23*, 3 (2013), 186–193. 67

[FW01] FITZPATRICK J. M., WEST J. B.: The distribution of target registration error in rigid-body point-based registration. *IEEE transactions on medical imaging 20*, 9 (2001), 917–927. 123, 136

[GBD04] GRAVEL P., BEAUDOIN G., DE GUISE, JACQUES A: A method for modeling noise in medical images. *IEEE transactions on medical imaging 23*, 10 (2004), 1221–1232. 81

[GBG*14] GERBER N., BELL B., GAVAGHAN K., WEISSTANNER C., CAVERSACCIO M., WEBER S.: Surgical planning tool for robotically assisted hearing aid implantation. *International journal of computer assisted radiology and surgery 9*, 1 (2014), 11–20. 2

[GBH70] GORDON R., BENDER R., HERMAN G. T.: Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of Theoretical Biology 29*, 3 (1970), 471–481. 67, 69, 70, 71, 105

[GBI09] GLASNER D., BAGON S., IRANI M.: Super-resolution from a single image. In *2009 IEEE 12th International Conference on Computer Vision (ICCV)* (2009), pp. 349–356. 17

[GEW10] GHOLIPOUR A., ESTROFF J. A., WARFIELD S. K.: Robust super-resolution volume reconstruction from slice acquisitions: application to fetal brain MRI. *IEEE transactions on medical imaging 29*, 10 (2010), 1739–1758. 35

[GHv*01] GOOS G., HARTMANIS J., VAN LEEUWEN J., NIESSEN W. J., VIERGEVER M. A. (Eds.): *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2001*. Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001. 131

[GKWT98] GUÉZIEC A., KAZANZIDES P., WILLIAMSON B., TAYLOR R. H.: Anatomy-based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot. *IEEE transactions on medical imaging 17*, 5 (1998), 715–728. 131

[GL14] GUILLEMOT C., LE MEUR O.: Image Inpainting: Overview and Recent Advances. *IEEE Signal Processing Magazine 31*, 1 (2014), 127–144. 30

[Glo81] GLOVER G. H.: An algorithm for the reduction of metal clip artifacts in CT reconstructions. *Medical Physics 8*, 6 (1981), 799. 116

[Gol07] GOLDMAN L. W.: Principles of CT: radiation dose and image quality. *Journal of nuclear medicine technology 35*, 4 (2007), 213–25; quiz 226–8. 28

[Gre08] GREENSPAN H.: Super-Resolution in Medical Imaging. *The Computer Journal 52*, 1 (2008), 43–63. 7

[GRL*98] GOLD S., RANGARAJAN A., LU C.-P., PAPPU S., MJOLSNESS E.: New algorithms for 2D and 3D point matching. *Pattern Recognition 31*, 8 (1998), 1019–1031. 152

[Had23] HADAMARD J.: *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. Yale University Press, New Haven, 1923. 10

[ham] HAMAMATSU: X-ray detectors. 64

[Han98] HANSEN P. C.: *Rank-deficient and discrete ill-posed problems: Numerical aspects of linear inversion*. SIAM monographs on mathematical modeling and computation. SIAM, Philadelphia, 1998. 10

[HBA97] HARDIE R. C., BARNARD K. J., ARMSTRONG E. E.: Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Transactions on Image Processing 6*, 12 (1997), 1621–1633. 11, 16

[HBB*06] HAMILL J. J., BRUNKEN R. C., BYBEL B., DIFILIPPO F. P., FAUL D. D.: A knowledge-based method for reducing attenuation artefacts caused by cardiac appliances in myocardial PET/CT. *Physics in medicine and biology 51*, 11 (2006), 2901–2918. 77

[HDIS09] HAMMING N. M., DALY M. J., IRISH J. C., SIEWERDSEN J. H.: Automatic image-to-world registration based on x-ray projections in cone-beam CT-guided interventions. *Medical physics 36*, 5 (2009), 1800–1812. 137

[Her09] HERMAN G. T.: *Fundamentals of computerized tomography: Image reconstruction from projections*, 2nd ed. ed. Advances in Pattern Recognition. Springer, Dordrecht and New York, 2009. 65

[Her11] HERBERT S.: Super-Resolution Reconstruction for Medical Image Enhancement using Uncertainty Models, 2011. 175

[HKK97] HONG M.-C., KANG M. G., KATSAGGELOS A. K.: Regularized multichannel restoration approach for globally optimal high-resolution video sequence. *Proc. SPIE 3024* (1997). 10

[HL94] HUDSON H. M., LARKIN R. S.: Accelerated image reconstruction using ordered subsets of projection data. *IEEE transactions on medical imaging 13*, 4 (1994), 601–609. 72, 76

[Hoe43] HOEL P. G.: On Indices of Dispersion. *The Annals of Mathematical Statistics 14*, 2 (1943), 155–162. 91

[Hou80] HOUNSFIELD G.: Computed medical imaging. *Science 210*, 4465 (1980), 22–28. 7

[HR94] H GUAN, R GORDON: A projection access order for speedy convergence of ART (algebraic reconstruction technique): a multilevel scheme for computed tomography. *Physics in medicine and biology 39*, 11 (1994), 2005. 72

[HYSI13] HIRAUMI H., YAMAMOTO N., SAKAMOTO T., ITO J.: A minimally invasive approach for cochlear implantation using a microendoscope. *European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery 270*, 2 (2013), 477–481. 2

[IP90] IRANI M., PELEG S.: Super resolution from image sequences. In *10th International Conference on Pattern Recognition* (16-21 June 1990), pp. 115–120. 11, 14

[IP91] IRANI M., PELEG S.: Improving Resolution by Image Registration. *CVGIP: Graph. Models Image Process. 53*, 3 (1991), 231–239. 11, 14, 18, 19

[ISM*] IGOR STENIN, STEFAN HANSEN, MARIA NAU-HERMES, WISSAM EL HAKIMI, MEIKE BECKER, JUDITH BEVERMANN, THOMAS KLENZNER, JOERG SCHIPPER: Evaluation von minimal invasiven multi-port Zugängen der Otobasis am humanen Schädelpräparat. In *Hubertus Feussner (Ed.) 2014 – 13. Jahrestagung der Deutschen Gesellschaft*. pp. 127–130.

[JDI*07] JAIN A., DEGUET A., IORDACHITA I., CHINTALAPANI G., BLEVINS J., LE Y., ARMOUR E., BURDETTE C., SONG D., FICHTINGER G.: Intra-operative 3D Guidance in Prostate Brachytherapy Using a Non-isocentric C-arm. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007*, Ayache N., Ourselin S., Maeder A., (Eds.), vol. 4792 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007, pp. 9–17. 134

[JF06] JAIN A., FICHTINGER G.: C-arm tracking and reconstruction without an external tracker. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention 9*, Pt 1 (2006), 494–502. 134

[JKKR13] JEONG K. Y., KIM J. H., KIM K. M., RA J. B.: Sinogram Super-Resolution Using a Space-Variant Blur Matrix in PET. *IEEE Transactions on Nuclear Science 60*, 1 (2013), 158–165. 18

[JMNM91] JACKSON J. I., MEYER C. H., NISHIMURA D. G., MACOVSKI A.: Selection of a convolution function for Fourier inversion using gridding computerised tomography application. *IEEE transactions on medical imaging 10*, 3 (1991), 473–478. 66

[JMZ*05] JAIN A. K., MUSTAFA T., ZHOU Y., BURDETTE C., CHIRIKJIAN G. S., FICHTINGER G.: FTRAC–a robust fluoroscope tracking fiducial. *Medical physics 32*, 10 (2005), 3185–3198. 134

[Jos82] JOSEPH, P. M.: An Improved Algorithm for Reprojecting Rays through Pixel Images. *IEEE transactions on medical imaging 1*, 3 (1982), 192–196. 69

[JSRF06] JANS H. S., SYME A. M., RATHEE S., FALLONE B. G.: 3D interfractional patient position verification using 2D-3D registration of orthogonal images. *Medical physics 33*, 5 (2006), 1420–1439. 131

[JSW81] JOSEPH, P. M., SCHULZ, R. A., WEISS G. H.: VIEW SAMPLING REQUIREMENTS IN FAN BEAM COMPUTED TOMOGRAPHY. *Journal of Computer Assisted Tomography 5*, 3 (1981). 57

[Kac37] KACZMARZ S.: Angenaherte auflosung Von systemen linearer gleichungenti bulletin international De lacademie polonaise des sciences et des Lettres. Classe des sciences mathematiques et naturelles. *Serie A, Sciences Mathematiques* (1937), 355–357. 69

[Kal11] KALENDER W.: *Computed tomography: Fundamentals, system technology, image quality, applications / by Willi A. Kalender*, 3rd rev. ed. ed. Wiley-VCH, Weinheim, 2011. 7, 8

[KAPS09] K. MOSALIGANTI, A. GELAS, P. COWGILL, S. MEGASON: An Optimized N-Dimensional Hough Filter for Detecting Spherical Image Objects. 137, 140

[KCWM12] KARIMI S., COSMAN P., WALD C., MARTZ H.: Segmentation of artifacts and anatomy in CT metal artifact reduction. *Medical physics 39*, 10 (2012), 5857–5868. 79, 116

[KHE87] KALENDER W. A., HEBEL R., EBERSBERGER J.: Reduction of CT artifacts caused by metallic implants. *Radiology 164*, 2 (1987), 576–577. 60, 78

[KHW13] KEYHANI M. H., HAKIMI W. E., WESARG S.: Anisotropy correction of medical image data employing patch similarity. In *2013 IEEE 26th International Symposium on Computer-Based Medical Systems (CBMS)* (2013), pp. 385–388. 18

[KIF*06] KENNEDY J. A., ISRAEL O., FRENKEL A., BAR-SHALOM R., AZHARI H.: Super-resolution in PET imaging. *IEEE transactions on medical imaging 25*, 2 (2006), 137–147. 18

[KIF*07] KENNEDY J. A., ISRAEL O., FRENKEL A., BAR-SHALOM R., AZHARI H.: The reduction of artifacts due to metal hip implants in CT-attenuation corrected PET images from hybrid PET/CT scanners. *Medical & biological engineering & computing 45*, 6 (2007), 553–562. 77

[KK07] KALENDER W. A., KYRIAKOU Y.: Flat-detector computed tomography (FD-CT). *European radiology 17*, 11 (2007), 2767–2779. 54

[KKK13] KOIVUKANGAS T., KATISKO J. P. A., KOIVUKANGAS J. P.: Technical accuracy of optical and the electromagnetic tracking systems. *SpringerPlus 2*, 1 (2013), 90. 157

[KNY*11] KONRAD P. E., NEIMAT J. S., YU H., KAO C. C., REMPLE M. S., D'HAESE P.-F., DAWANT B. M.: Customized, miniature rapid-prototype stereotactic frames for use in deep brain stimulator surgery: initial clinical methodology and experience from 263 patients from 2002 to 2008. *Stereotactic and functional neurosurgery 89*, 1 (2011), 34–41. 125

[KP09] KONG H., PAN J.: An Improved Ordered-Subset Simultaneous Algebraic Reconstruction Technique. In *2009 2nd International Congress on Image and Signal Processing (CISP)* (2009), pp. 1–5. 67

[KPB88] KEREN D., PELEG S., BRADA R.: Image sequence enhancement using sub-pixel displacements. In *Computer Vision and Pattern Recognition, 1988. Proceedings CVPR '88., Computer Society Conference on* (1988), pp. 742–746. 11, 14

[KR02] KASAP S. O., ROWLANDS J. A.: Direct-conversion flat-panel X-ray image detectors. *IEE Proceedings - Circuits, Devices and Systems 149*, 2 (2002), 85–96. 55

[KS93] KIM S. P., SU W. Y.: Recursive high-resolution reconstruction of blurred multiframe images. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society 2*, 4 (1993), 534–539. 12

[KS01] KAK A. C., SLANEY M.: *Principles of computerized tomographic imaging*, vol. 33 of *Classics in applied mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, 2001. 65, 67, 72, 79, 105

[La 05] LA RIVIERE, PATRICK J: Penalized-likelihood sinogram smoothing for low-dose CT. *Medical physics 32*, 6 (2005), 1676–1683. 75

[LAH03] LASCALZA S., ARICO J., HUGHES R.: Effect of metal and sampling rate on accuracy of Flock of Birds electromagnetic tracking system. *Journal of Biomechanics 36*, 1 (2003), 141–144. 157

[LBKH00] LAROSE D., BAYOUTH J., KANADE T., HANSON K. M.: Transgraph: interactive intensity-based 2D/3D registration of x-ray and CT data. In *Medical Imaging 2000* (2000), SPIE Proceedings, SPIE, pp. 385–396. 131

[LBN*14] LABADIE R. F., BALACHANDRAN R., NOBLE J. H., BLACHON G. S., MITCHELL J. E., REDA F. A., DAWANT B. M., FITZPATRICK J. M.: Minimally invasive image-guided cochlear implantation surgery: first report of clinical implementation. *The Laryngoscope 124*, 8 (2014), 1915–1922. 166, 167

[LD04] LEO JOSKOWICZ, DOTAN KNAAN: How to achieve fast, accurate, and robust rigid registration between fluoroscopic X-ray and CT images of the 18th International Congress and Exhibition, Chicago, USA, June 23-26, 2004. In *CARS 2004. Computer Assisted Radiology and Surgery. Proceedings of the 18th International Congress and Exhibition, Chicago, USA, June 23-26, 2004* (2004), Heinz U. Lemke, Kiyonari Inamura, Kunio Doi, Michael W. Vannier, Allan G. Farman, Johan H. C. Reiber, (Eds.), vol. 1268 of *International Congress Series*, Elsevier, pp. 147–152. 165, 166

[LF15] LEMMERLING M., FOER B. D.: *Temporal bone imaging*. Medical radiology. Diagnostic imaging. Springer, Berlin, 2015. 9

[LHWG12] LIEBERMAN I. H., HARDENBROOK M. A., WANG J. C., GUYER R. D.: Assessment of pedicle screw placement accuracy, procedure time, and radiation exposure using a miniature robotic guidance system. *Journal of spinal disorders & techniques 25*, 5 (2012), 241–248. 165, 166

[LL93] LI C. H., LEE C. K.: Minimum cross entropy thresholding. *Pattern Recognition 26*, 4 (1993), 617–625. 141

[LMK*12] LELL M. M., MEYER E., KUEFNER M. A., MAY M. S., RAUPACH R., UDER M., KACHELRIESS M.: Normalized metal artifact reduction in head and neck computed tomography. *Investigative radiology 47*, 7 (2012), 415–421. 78, 116

[LND*08] LABADIE R. F., NOBLE J. H., DAWANT B. M., BALACHANDRAN R., MAJDANI O., FITZPATRICK J. M.: Clinical validation of percutaneous cochlear implant surgery: initial report. *The Laryngoscope 118*, 6 (2008), 1031–1039. 2

[LOA*14] LIU W. P., OTAKE Y., AZIZIAN M., WAGNER O. J., SORGER J. M., ARMAND M., TAYLOR R. H.: 2D-3D radiograph to cone-beam computed tomography (CBCT) registration for C-arm image-guided robotic surgery. *International journal of computer assisted radiology and surgery* (2014). 165, 166

[LXQ13] LEI WANG, XIN GAO, QIANG FANG: A novel mutual information-based similarity measure for 2D/3D registration in image guided intervention. In *2013 1st International Conference on Orange Technologies (ICOT 2013)* (2013), pp. 135–138. 131

[LYJ03] LIVYATAN H., YANIV Z., JOSKOWICZ L.: Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. *IEEE transactions on medical imaging 22*, 11 (2003), 1395–1406. 132

[LYM*10] LIU X., YU L., MANDUCA A., RITMAN E. L., MCCOLLOUGH C. H.: A super resolution technique for clinical multislice CT. In *SPIE Medical Imaging* (2010), Samei E., Pelc N. J., (Eds.), SPIE Proceedings, SPIE, pp. 76221Q–76221Q–7. 18

[MDNK10] MAASS C., DENNERLEIN F., NOO F., KACHELRIESS M.: Comparing short scan CT reconstruction algorithms regarding cone-beam artifact performance. In *2010 IEEE Nuclear Science Symposium and Medical Imaging Conference (2010 NSS/MIC)* (2010), pp. 2188–2193. 67

[Meh14] MEHLTRETTER M.: Optimierung einer 2D/3D Registrierungsmethodik für bildgestützte Eingriffe, 2014. 175

[MGK*95] MANGLOS S. H., GAGNE G. M., KROL A., THOMAS F. D., NARAYANASWAMY R.: Transmission maximum-likelihood reconstruction with ordered subsets for cone beam CT. *Physics in medicine and biology 40*, 7 (1995), 1225–1241. 68, 75, 76

[MNS00] MITSCHKE M. M., NAVAB N., SCHUETZ O.: Online geometrical calibration of a mobile C-arm using external sensors. In *Medical Imaging 2000* (2000), Mun S. K., (Ed.), SPIE Proceedings, SPIE, pp. 580–587. 82

[MRA*06] MIQUEL M. E., RHODE K. S., ACHER P. L., MACDOUGALL N. D., BLACKALL J., GASTON R. P., HEGDE S., MORRIS S. L., BEANEY R., DEEHAN C., POPERT R., KEEVIL S. F.: Using combined x-ray and MR imaging for prostate I-125 post-implant dosimetry: phantom validation and preliminary patient work. *Physics in medicine and biology 51*, 5 (2006), 1129–1137. 130

[MRL*10] MEYER E., RAUPACH R., LELL M., SCHMIDT B., KACHELRIESS M.: Normalized metal artifact reduction (NMAR) in computed tomography. *Medical physics 37*, 10 (2010), 5482–5493. 78, 79

[MSC*11] M. DIRK R., STEPHANIE J. C., CYNTHIA A. T., JOSEPH A. I., LO JOSEPH Y., SINA F.: *New Applications of Super-resolution in Medical Imaging*. Super-resolution imaging. CRC Press, Boca Raton, 2011. 18

[NBHM*96] NAVAB N., BANI-HASHEMI A. R., MITSCHKE M. M., HOLDSWORTH D. W., FAHRIG R., FOX A. J., GRAUMANN R.: Dynamic geometrical calibration for 3D cerebral angiography. In *Medical Imaging 1996* (1996), van Metter R. L., Beutel J., (Eds.), SPIE Proceedings, SPIE, pp. 361–370. 133

[NBHN*98] NAVAB N., BANI-HASHEMI A., NADAR M. S., WIESENT K., DURLAK P., BRUNNER T., BARTH K., GRAUMANN R.: 3D reconstruction from projection matrices in a C-arm based 3D-angiography system. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*, Goos G., Hartmanis J., van Leeuwen J., Wells W. M., Colchester A., Delp S., (Eds.), vol. 1496 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1998, pp. 119–129. 163

[NDMS05] NICHOLAS KEAT, DAVID PLATTEN, MARIA LEWIS, SUE EDYVEAN: Report 05071, Siemens Somatom Sensation open CT scanner technical evaluation. 29

[NHPS14] NAU-HERMES M., POLLMANNS S., SCHMITT R.: Uncertainty Evaluation for Surgical Processes. *Key Engineering Materials 613* (2014), 317–326. 159

[NHSB*14] NAU-HERMES M., SCHMITT R., BECKER M., EL-HAKIMI W., HANSEN S., KLENZNER T., SCHIPPER J.: Quality assurance of multiport image-guided minimally invasive surgery at the lateral skull base. *BioMed research international 2014* (2014), 904803. 169

[NM00] NGUYEN N., MILANFAR P.: An efficient wavelet-based algorithm for image superresolution. In *Proceedings of 7th IEEE International Conference on Image Processing* (2000), pp. 351–354 vol.2. 13

[NPS13] NAU M., POLLMANNS S., SCHMITT R.: Assessing the risk of minimally-invasive surgery: a metrological approach. In *16th International Congress of Metrology* (October 7-10, 2013), Filtz J.-R., Larquier B., Claudel P., Favreau J.-O., (Eds.), p. 7002. v, 125

[NTC04] NING R., TANG X., CONOVER D.: X-ray scatter correction algorithm for cone beam CT imaging. *Medical Physics 31*, 5 (2004), 1195. 58

[OB09] OEHLER M., BUZUG T. M.: An Image Inpainting Based Surrogate Data Strategy for Metal Artifact Reduction in CT Images. In *4th European Conference of the International Federation for Medical and Biological Engineering*, Vander Sloten J., Verdonck P., Nyssen M., Haueisen J., (Eds.), vol. 22 of *IFMBE Proceedings*. Springer Berlin Heidelberg, 2009, pp. 651–654. 117

[O'S85] O'SULLIVAN J. D.: A fast sinc function gridding algorithm for fourier inversion in computer tomography. *IEEE transactions on medical imaging 4*, 4 (1985), 200–207. 66

[OSB99] OPPENHEIM A. V., SCHAFER R. W., BUCK J. R.: *Discrete-time signal processing*, 2nd ed ed. Prentice Hall signal processing series. Prentice Hall, Upper Saddle River, N.J., 1999. 66

[Par82] PARKER D. L.: Optimal short scan convolution reconstruction for fanbeam CT. *Medical physics 9*, 2 (1982), 254–257. 66

[PDBF12] P. MARKELJ, D. TOMAZEVIC, B. LIKAR, F. PERNUS: A review of 3D/2D registration methods for image-guided interventions. *Medical Image Analysis 16*, 3 (2012), 642–661. 37, 123, 129, 130, 132, 138

[PEH*07] PENNEY G. P., EDWARDS P. J., HIPWELL J. H., SLOMCZYKOWSKI M., REVIE I., HAWKES D. J.: Postoperative calculation of acetabular cup position using 2-D-3-D registration. *IEEE transactions on bio-medical engineering 54*, 7 (2007), 1342–1348. 131

[PJH*06] PAUL A. YUSHKEVICH, JOSEPH PIVEN, HEATHER CODY HAZLETT, RACHEL GIMPEL SMITH, SEAN HO, JAMES C. GEE, GUIDO GERIG: User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage 31*, 3 (2006), 1116–1128. 44

[PKN*04] PEETERS R. R., KORNPROBST P., NIKOLOVA M., SUNAERT S., VIEVILLE T., MALANDAIN G., DERICHE R., FAUGERAS O., NG M., VAN HECKE P.: The use of super-resolution techniques to reduce slice thickness in functional MRI. *International Journal of Imaging Systems and Technology 14*, 3 (2004), 131–138. 17

[PKS87] PELEG S., KEREN D., SCHWEITZER L.: Improving image resolution using subpixel motion. *Pattern Recognition Letters 5*, 3 (1987), 223–226. 11, 14

[Ple13] PLENGE E.: *Super-Resolution and Self-Similarity in Magnetic Resonance Imaging: Super-resolutie en zelf-similariteit in magnetische resonantie beeldvorming = Super-resolutie en zelf-similariteit in magnetische resonantie beeldvorming*. Erasmus Universiteit Rotterdam, Rotterdam, 2013. 19

[Pow64] POWELL, M. J. D.: An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The Computer Journal 7*, 2 (1964), 155–162. 86

[PRd10] PAPAL C., RONGEN P. M. J., DE WITH,PETER H. N.: Surgical needle reconstruction using small-angle multi-view X-ray. In *2010 17th IEEE International Conference on Image Processing (ICIP 2010)* (2010), pp. 4193–4196. 158

[PWL*98] PENNEY G. P., WEESE J., LITTLE J. A., DESMEDT P., HILL D. L. G., HAWKES D. J.: A comparison of similarity measures for use in 2D-3D medical image registration. In *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*, Goos G., Hartmanis J., van Leeuwen J., Wells W. M., Colchester A., Delp S., (Eds.), vol. 1496 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1998, pp. 1153–1161. 131

[PY01] PELED S., YESHURUN Y.: Superresolution in MRI: application to human white matter fiber tract visualization by diffusion tensor imaging. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 45*, 1 (2001), 29–35. 17

[RGN13] REICHL T., GARDIAZABAL J., NAVAB N.: Electromagnetic servoing-a new tracking paradigm. *IEEE transactions on medical imaging 32*, 8 (2013), 1526–1535. 157

[RKS10a] ROUSSEAU F., KIM K., STUDHOLME C.: A groupwise super-resolution approach: Application to brain MRI. In *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (2010), pp. 860–863. 17

[RKS*10b] ROUSSEAU F., KIM K., STUDHOLME C., KOOB M., DIETEMANN J. L.: On super-resolution for fetal brain MRI. *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention 13*, Pt 2 (2010), 355–362. 35

[Rön96] RÖNTGEN W. C.: *Ueber eine neue Art von Strahlen*, 1 ed. Sitzungsberichte der Würzburger Physik.-medic. Gesellschaft. Stahel'sche K. Hof- und Universitätsbuch- und Kunsthandlung, Würzburg, 1896. 7

[ROP*13] ROUSSEAU F., OUBEL E., PONTABRY J., SCHWEITZER M., STUDHOLME C., KOOB M., DIETEMANN J.-L.: BTK: an open-source toolkit for fetal brain MR image processing. *Computer methods and programs in biomedicine 109*, 1 (2013), 65–73. 35

[RR96] RICHARD R. SCHULTZ, ROBERT L. STEVENSON: Extraction of High-Resolution Frames from Video Sequences. *IEEE Transactions on Image Processing 5* (1996), 996–1011. 11, 16, 17

[RSM12] ROBERT SCHMITT, SEBASTIAN POLLMANNS, MARIA NAU: Accuracy Study of medical CT Systems for Image-Guided Surgery according to VDI/VDE GMA 2630 Guideline. In *Conference on Industrial Computed Tomography (ICT)*. 2012, pp. 209–217. 37

[RW10] RAHMAN S. U., WESARG S.: Upsampling of cardiac MR images: Comparison of averaging and super-resolution for the combination of multiple views. In *2010 10th IEEE International Conference on Information Technology and Applications in Biomedicine (ITAB 2010)* (2010), pp. 1–4. 17

[Sak93] SAKAS G.: *Fraktale Wolken, virtuelle Flammen: Computer-Emulation und Visualisierung turbulenter Gasbewegung*. Beiträge zur graphischen Datenverarbeitung. Springer-Verlag, Berlin and New York, 1993. 70

[SCL*13] SHI W., CABALLERO J., LEDIG C., ZHUANG X., BAI W., BHATIA K., DE MARVAO,ANTONIO M SIMOES MONTEIRO, DAWES T., O'REGAN D., RUECKERT D.: Cardiac Image Super-Resolution with Global Correspondence Using Multi-Atlas PatchMatch. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, Hutchison D., Kanade T., Kittler J., Kleinberg J. M., Mattern F., Mitchell J. C., Naor M., Nierstrasz O., Pandu Rangan C., Steffen B., Sudan M., Terzopoulos D., Tygar D., Vardi M. Y., Weikum G., Mori K., Sakuma I., Sato Y., Barillot C., Navab N., (Eds.), vol. 8151 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 9–16. 17

[SDW14] STEGER T., DRECHSLER K., WESARG S.: Image-Based Bronchoscopy Navigation System Based on CT and C-arm Fluoroscopy. In *Clinical Image-Based Procedures. Translational Research in Medical Imaging*, Erdt M., Linguraru M. G., Oyarzun Laura C., Shekhar R., Wesarg S., GonzÃ¡lez Ballester M. A., Drechsler K., (Eds.), vol. 8361 of *Lecture Notes in Computer Science*. Springer International Publishing, Cham, 2014, pp. 15–22. 135, 136

[Seb14] SEBASTIAN POLLMANS: *Bestimmung von Unsicherheitsbeiträgen bei medizinischen Computertomografiemessungen für die bildbasierte navigierte Chirurgie*. PhD thesis, Fakultät für Maschinenwesen, Rheinisch-Westfälischen Technischen Hochschule Aachen, 2014. 36

[SEE11] STEGER S., EBERT D., ERDT M.: Lymph node segmentation in CT slices using dynamic programming. In *8th IEEE International Symposium on Biomedical Imaging (ISBI 2011)* (2011), pp. 1990–1993. 138, 154

[SEHW14] STEGER T., EL-HAKIMI W., WESARG S.: Tracking von Instrumenten auf fluoroskopischen Aufnahmen für die navigierte Bronchoskopie. In *Bildverarbeitung für die Medizin 2014*, Deserno T. M., Handels H., Meinzer H.-P., Tolxdorff T., (Eds.), Informatik aktuell. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 228–233.

[SH12] STEGER T., HOSBACH M.: Navigated bronchoscopy using intraoperative fluoroscopy and preoperative CT. In *2012 IEEE 9th International Symposium on Biomedical Imaging (ISBI 2012)* (2012), pp. 1220–1223. 82, 135

[SHG*11] SCHULZE R., HEIL U., GROSS D., BRUELLMANN D. D., DRANISCHNIKOW E., SCHWANECKE U., SCHOEMER E.: Artefacts in CBCT: a review. *Dento maxillo facial radiology 40*, 5 (2011), 265–273. 60, 67, 81

[Sid85] SIDDON R. L.: Fast calculation of the exact radiological path for a three-dimensional CT array. *Medical physics 12*, 2 (1985), 252–255. 69, 70

[Sie] SIEMENS AG: SOMATOM Definition Flash. 9, 125

[Sie11] SIEWERDSEN J. H.: Cone-Beam CT with a Flat-Panel Detector: From Image Science to Image-Guided Surgery. *Nuclear instruments & methods in physics research. Section A, Accelerators, spectrometers, detectors and associated equipment 648*, S1 (2011), S241–S250. 123, 132

[Sie12] SIEMENS AG: Image-SOMATOM-Emotion, 17.02.2012. 9

[SJ01] SIEWERDSEN J. H., JAFFRAY D. A.: Cone-beam computed tomography with a flat-panel imager: magnitude and effects of x-ray scatter. *Medical physics 28*, 2 (2001), 220–231. 58

[SJA10] SAUERLAND S., JASCHINSKI T., AM NEUGEBAUER E.: Laparoscopic versus open surgery for suspected appendicitis. *The Cochrane database of systematic reviews*, 10 (2010), CD001546. 1

[SL74] SHEPP L. A., LOGAN B. F.: The Fourier reconstruction of a head section. *IEEE Transactions on Nuclear Science 21*, 3 (1974), 21–43. 70, 108

[SMB*04] SIEWERDSEN J. H., MOSELEY D. J., BAKHTIAR B., RICHARD S., JAFFRAY D. A.: The influence of antiscatter grids on soft-tissue detectability in cone-beam computed tomography with flat-panel detectors. *Medical Physics 31*, 12 (2004), 3506. 58

[SMW*09] SIEWERDSEN J. H., MIGA M. I., WONG K. H., DALY M. J., CHAN H., NITHIANANTHAN S., HAMMING N., BROCK K. K., IRISH J. C.: High-performance intraoperative cone-beam CT on a mobile C-arm: an integrated system for guidance of head and neck surgery. In *SPIE Medical Imaging* (2009), SPIE Proceedings, SPIE, p. 72610J. 157

[SO89] STARK H., OSKOUI P.: High-resolution image recovery from image-plane arrays, using convex projections. *Journal of the Optical Society of America. A, Optics and image science 6*, 11 (1989), 1715–1726. 11, 16

[SSZ*12] SCHAFER S., STAYMAN J. W., ZBIJEWSKI W., SCHMIDGUNST C., KLEINSZIG G., SIEWERDSEN J. H.: Antiscatter grids in mobile C-arm cone-beam CT: effect on image quality and dose. *Medical physics 39*, 1 (2012), 153–159. 81

[ST88] SAMET H., TAMMINEN M.: Efficient component labeling of images of arbitrary dimension represented by linear bintrees. *IEEE Transactions on Pattern Analysis and Machine Intelligence 10*, 4 (1988), 579–586. 139, 141

[Ste14] STEGER T.: *Bronchoskopische Navigation mittels Pose Estimation des C-Bogens aus musterkodierten Fluoroskopie-Aufnahmen*. PhD thesis, Darmstadt, TU, Diss., 2014, 2014. 86, 129, 137, 152

[SV82] SHEPP L. A., VARDI Y.: Maximum likelihood reconstruction for emission tomography. *IEEE transactions on medical imaging 1*, 2 (1982), 113–122. 75

[SZB*13] SISNIEGA A., ZBIJEWSKI W., BADAL A., KYPRIANOU I. S., STAYMAN J. W., VAQUERO J. J., SIEWERDSEN J. H.: Monte Carlo study of the effects of system geometry and antiscatter grids on cone-beam CT scatter distributions. *Medical physics 40*, 5 (2013), 51915. 58

[TA77] TIKHONOV A. N., ARSENIN V. Y.: *Solutions of Ill-posed problems*. W.H. Winston, 1977. 10, 12

[TA98] TAGUCHI K., ARADATE H.: Algorithm for image reconstruction in multi-slice helical CT. *Medical physics 25*, 4 (1998), 550–561. 28

[TH84] TSAI R., HUANG T.: Multiframe image restoration and registration. In *Advances in Computer Vision and Image Processing* (1984). 12

[TK94] TOM B. C., KATSAGGELOS A. K.: Multichannel image identification and restoration using the expectation-maximization algorithm. *Proc. SPIE 2298* (1994). 11, 16

[TLSP03] TOMAZEVIC D., LIKAR B., SLIVNIK T., PERNUS F.: 3-D/2-D registration of CT and MR to X-ray images. *IEEE transactions on medical imaging 22*, 11 (2003), 1407–1416. 132

[TM11] TIAN J., MA K.-K.: A survey on super-resolution imaging. *Signal, Image and Video Processing 5*, 3 (2011), 329–342. 12, 17

[TOS92] TEKALP A., OZKAN M. K., SEZAN M.: High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration. In *Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992 IEEE International Conference on* (1992), vol. 3, pp. 169–172 vol.3. 11, 12, 16

[Tuy83] TUY H. K.: An Inversion Formula for Cone-Beam Reconstruction. *SIAM Journal on Applied Mathematics 43*, 3 (1983), 546–552. 57, 65

[UAE91] UNSER M., ALDROUBI A., EDEN M.: Fast B-spline transforms for continuous image representation and interpolation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 13*, 3 (1991), 277–285. 13

[UG92] UR H., GROSS D.: Improved Resolution from Subpixel Shifted Pictures. *CVGIP: Graph. Models Image Process. 54*, 2 (1992), 181–186. 13

[vDH*11] VAN DER BOM,I. M. J., DAWANT B. M., HAYNOR D. R., KLEIN S., STARING M., HOMAN R., BARTELS L. W., PLUIM J. P. W.: Evaluation of optimization methods for intensity-based 2D-3D registration in x-ray guided interventions. In *SPIE Medical Imaging* (2011), SPIE Pro-

ceedings, SPIE, p. 796223. 86, 152

[vEPR02] VAN ASSEN, H C, EGMONT-PETERSEN M., REIBER, J H C: Accurate object localization in gray level images using the center of gravity measure: accuracy versus precision. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society 11*, 12 (2002), 1379–1384. 137

[VJvG10] VELDKAMP, WOUTER J H, JOEMAI, RAOUL M S, VAN DER MOLEN, AART J, GELEIJNS J.: Development and validation of segmentation and interpolation techniques in sinograms for metal artifact suppression in CT. *Medical physics 37*, 2 (2010), 620–628. 78

[WBE11] WESARG S., BECKER M., EL HAKIMI W.: Planning of High Precision Surgery at the Lateral Skull Base. In *14th Korea-Germany Joint Workshop on Advanced Medical Image Processing 2011. Heidelberg, Germany*. 2011, pp. 72–78. 8

[WBM*09] WANNA G., BALACHANDRAN R., MAJDANI O., MITCHELL J., LABADIE R.: Percutaneous access to the petrous apex in vitro using customized micro-stereotactic frames based on image-guided surgical technology. *Acta Oto-Laryngologica* (2009), 1–6. 2

[Web90] WEBB S.: *From the watching of shadows: The origins of radiological tomography*. 1990. 7

[Wei97] WEICKERT J.: A review of nonlinear diffusion filtering. In *Scale-Space Theory in Computer Vision*, Goos G., Hartmanis J., Leeuwen J., Haar Romeny B., Florack L., Koenderink J., Viergever M., (Eds.), vol. 1252 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1997, pp. 1–28. 31

[WLB*98] WANG L., LAI H.-M., BARKER G. J., MILLER D. H., TOFTS P. S.: Correction for variations in MRI scanner sensitivity in brain studies with histogram matching. *Magnetic Resonance in Medicine 39*, 2 (1998), 322–327. 48

[WLL14] WOLF D., LUBK A., LICHTE H.: Weighted simultaneous iterative reconstruction technique for single-axis tomography. *Ultramicroscopy 136* (2014), 15–25. 67

[WLZ*13] WANG Q., LI L., ZHANG L., CHEN Z., KANG K.: A novel metal artifact reducing method for cone-beam CT based on three approximately orthogonal projections. *Physics in medicine and biology 58*, 1 (2013), 1–17. 78

[WME*06] WHITING B. R., MASSOUMZADEH P., EARL O. A., O'SULLIVAN J. A., SNYDER D. L., WILLIAMSON J. F.: Properties of preprocessed sinogram data in x-ray computed tomography. *Medical physics 33*, 9 (2006), 3290–3303. 81

[WMFM96] WANG M. Y., MAURER, C R JR, FITZPATRICK J. M., MACIUNAS R. J.: An automatic technique for finding and localizing externally attached markers in CT and MR volume images of the head. *IEEE transactions on bio-medical engineering 43*, 6 (1996), 627–637. 140

[WXZ14] WU X., XIAO S., ZHANG Y.: Registration based super-resolution reconstruction for lung 4D-CT. *Conference proceedings : ... Annual International Conference of the IEEE Engineering*

in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference 2014 (2014), 2444–2447. 18

[YCDS12] YATZIV L., CHARTOUNI M., DATTA S., SAPIRO G.: Toward multiple catheters detection in fluoroscopic image guided interventions. *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society 16*, 4 (2012), 770–781. 158

[YCT*12] YANG C.-D., CHEN Y.-W., TSENG C.-S., HO H.-J., WU C.-C., WANG K.-W.: Non-invasive, fluoroscopy-based, image-guided surgery reduces radiation exposure for vertebral compression fractures: A preliminary survey. *Formosan Journal of Surgery 45*, 1 (2012), 12–19. 165, 166

[YGP*12] YU ZHANG, GUORONG WU, PEW-THIAN YAP, QIANJIN FENG, JUN LIAN, WUFAN CHEN, DINGGANG SHEN: Reconstruction of super-resolution lung 4D-CT using patch-based sparse representation. In *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2012), pp. 925–931. 18

[YJ05] YANIV Z., JOSKOWICZ L.: Precise robot-assisted guide positioning for distal locking of intramedullary nails. *IEEE transactions on medical imaging 24*, 5 (2005), 624–635. 158

[YLL*09] YU L., LIU X., LENG S., KOFLER J. M., RAMIREZ-GIRALDO J. C., QU M., CHRISTNER J., FLETCHER J. G., MCCOLLOUGH C. H.: Radiation dose reduction in computed tomography: techniques and future perspective. *Imaging in medicine 1*, 1 (2009), 65–84. 58, 75

[YMW09] YANIV Z., MIGA M. I., WONG K. H.: Fiducial localization in C-arm based cone-beam CT. In *SPIE Medical Imaging* (2009), SPIE Proceedings, SPIE, pp. 72610I–72610I–12. 137

[YTZ*09] YANG P., TONG X., ZHENG X., ZHENG J., HE Y.: A Gradient-Based Adaptive Interpolation Filter for Multiple View Synthesis. In *Advances in Multimedia Information Processing - PCM 2009*, Muneesawang P., Wu F., Kumazawa I., Roeksabutr A., Liao M., Tang X., (Eds.), vol. 5879 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2009, pp. 551–560. 30, 31

[ZGS14] ZHAO Z., GANG G. J., SIEWERDSEN J. H.: Noise, sampling, and the number of projections in cone-beam CT with a flat-panel detector. *Medical physics 41*, 6 (2014), 061909. 57

[ZPH*11] ZHANG Y., PU Y.-F., HU J.-R., LIU Y., ZHOU J.-L.: A new CT metal artifacts reduction algorithm based on fractional-order sinogram inpainting. *Journal of X-ray science and technology 19*, 3 (2011), 373–384. 79

[ZRW*00] ZHAO S., ROBERTSON D. D., WANG G., WHITING B., BAE K. T.: X-ray CT metal artifact reduction using wavelets: an application for imaging total hip prostheses. *IEEE transactions on medical imaging 19*, 12 (2000), 1238–1247. 77